

DISPLAY SYSTEM FOR DOCUMENT, RECEIVING APPARATUS, DISPLAY METHOD THEREFOR AND RECORDING MEDIUM

Publication number: JP2003108278 (A)

Publication date: 2003-04-11

Inventor(s): KASAI YASUO; YAMADA KOJI; NAGAHO KAZUYOSHI; SUZUKI KATSUHITO +

Applicant(s): SEIKO EPSON CORP +

Classification:

- international: G06F3/00; G06F3/048; G06F3/00; G06F3/048; (IPC1-7): G06F3/00

- European:

Application number: JP20010303205 20010928

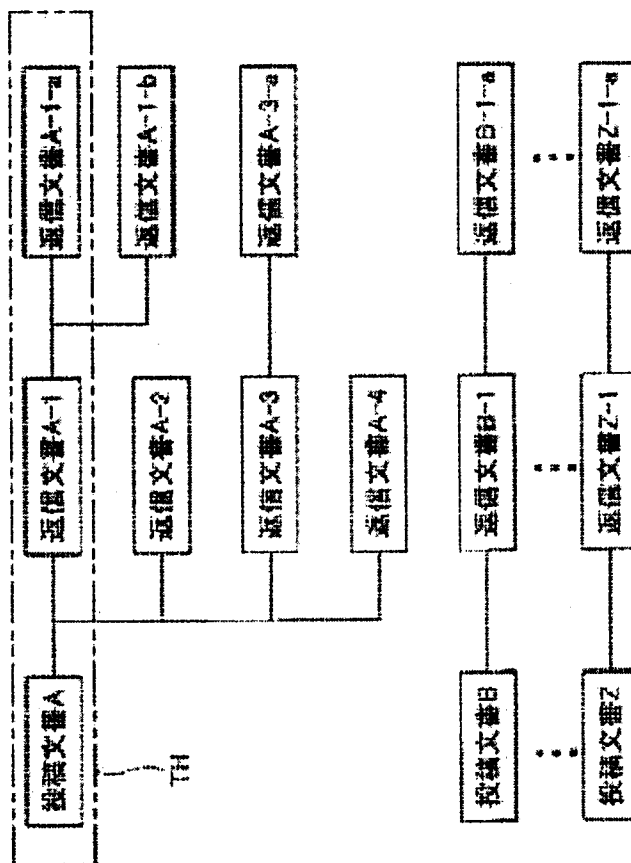
Priority number(s): JP20010303205 20010928

Also published as:

JP4058928 (B2)

Abstract of JP 2003108278 (A)

PROBLEM TO BE SOLVED: To allow a series of documents having sending-and-replying relationships to be understood for outlines of each document quickly and clearly and select document information easily and efficiently. SOLUTION: In a display system for bulletin board KS, each thread TH of an electronic bulletin board extracts important sentences from each document (submitted documents A, replying documents A-1 and replying documents A-1-a) belonging to each thread TH, and creates abstracts of each thread from each important sentence to display the abstracts on a screen. When extracting the important sentences in consideration of contents represented by body copy of the whole documents belonging to a single thread TH, sentences with the nearest meaning to the contents represented by the body copy are extracted from body copy of each document as the important sentences.



* NOTICES *

JPFO and INPIT are not responsible for any damages caused by the use of this translation.

- 1.This document has been translated by computer. So the translation may not reflect the original precisely.
- 2.**** shows the word which can not be translated.
- 3.In the drawings, any words are not translated.

CLAIMS

[Claim(s)]

[Claim 1]A document display system which displays on a screen at least some two or more documents which have a predetermined relation so that an inspection is possible, comprising:
A document group extraction means to extract a mass of document group from said two or more documents.

A significant sentence extraction means to extract a sentence which has predetermined correlation in this document as a significant sentence from each document belonging to said document group in consideration of contents expressed by the whole document belonging to said document group.

A significant sentence displaying means which displays a significant sentence extracted by this significant sentence extraction means on said screen as some documents belonging to said document group.

[Claim 2]A document display system which is a document which two or more documents which are the document display systems according to claim 1, and have said predetermined relation have in relation between a letter sent to get a reply and a reply.

[Claim 3]A document display system which is the document display system according to claim 1 or 2, and is a means to extract said significant sentence according to a kind which was provided with a kind discriminating means which distinguishes a kind of said document group, and from which said significant sentence extraction means was distinguished by this kind discriminating means.

[Claim 4]The document display system according to claim 3 which is at least one of asked type of a question with which a kind of said document group includes question expression, two or more subject type including two or more subjects, and interactive modes containing conversational sentences.

[Claim 5]The document display system according to claim 3 or 4 provided with a correlation setting-out means to change a standard which extracts a sentence which has predetermined correlation in this document as a significant sentence according to a kind of document belonging to this document group, from each document in which said significant sentence extraction means belongs to said document group.

[Claim 6]Are the document display system according to any one of claims 1 to 5, and it has a summary preparing means which creates a summary of said document group based on a significant sentence extracted by said significant sentence extraction means, A document display system provided with a summary displaying means which displays a summary which changed to said significant sentence displaying means, and was created by this summary preparing means on said screen.

[Claim 7]The document display system according to claim 6 provided with a significant sentence displaying means which displays a significant sentence extracted by said significant sentence extraction means on said screen as some documents belonging to said document group.

[Claim 8]A document display system provided with a sentence standardization means which is the document display system according to any one of claims 1 to 7, and rewrites a sentence of each of said document to a standard expression based on a predetermined standard.

[Claim 9]A document display system which is the document display stem according to claim 6 or 7, and is provided with a means which rewrites expression of a significant sentence said summary preparing means was extracted by said significant sentence extraction means at least to a standard expression.

[Claim 10]It has an owner meaning sentence extraction means to be the document display system according to any one of claims 1 to 9, and to extract a sentence except expression which cannot constitute a summary of a document out of said document based on a predetermined standard as an owner meaning sentence, A document display system which is a means by which said significant sentence extraction means or said summary preparing means performs extraction of said significant sentence, or creation of said summary using this owner meaning sentence.

[Claim 11]The document display system according to any one of claims 1 to 10 provided with a means by which said significant sentence displaying means or a summary displaying means carries out the list display of a significant sentence about said document group, or the summary to the order of a time series.

[Claim 12]The document display system according to claim 11 provided with a means for said significant sentence displaying means or a summary displaying means to relate a maker of a document belonging to said document group with said significant sentence or said summary, and to display.

[Claim 13]It is the document display system according to any one of claims 1 to 12, A document display system which was provided with a title preparing means which creates a title of said document group based on extraction by said significant sentence extraction means, and was provided with a means for said significant sentence displaying means or a summary displaying means to relate with said significant sentence or said summary a title of each document drawn up by this title preparing means, and to display.

[Claim 14]A receiving set which receives a document via a communication line, comprising: A document group extraction means to extract a mass of document group from said two or more received documents.

This extraction means which takes out the text from each document belonging to this document group at least about a document group of 1 extracted by this document group extraction means.

An extraction means to extract a sentence which has predetermined correlation in this document as a significant sentence from each document belonging to said document group in consideration of contents expressed by the whole text taken out by this this extraction means.

An output means which outputs a significant sentence about said document group of 1 extracted by this extraction means via a communication line.

[Claim 15]A receiving set which is the receiving set according to claim 14, was provided with a preparing means which creates a summary of said document group based on a significant sentence extracted by said extraction means, and was provided with a summary output means which outputs a summary which changed to said output means and was created by this preparing means via a communication line.

[Claim 16]Are a document display method which displays on a screen at least some two or more documents which have a predetermined relation so that an inspection is possible, extract a mass of document group from said two or more documents, and contents expressed by the whole document belonging to said document group are taken into consideration, A document display method which extracts a sentence which has predetermined correlation in this document as a significant sentence from each document belonging to said document group, and displays a this extracted significant sentence on said screen as some documents belonging to said document group.

[Claim 17]A process of being the recording medium which recorded a computer program for displaying on a screen at least some two or more documents which have a predetermined relation so that an inspection is possible, and extracting a mass of document group from said two or more documents, A process of extracting a sentence which has predetermined correlation in this document as a significant sentence from each document belonging to said document group in consideration of contents expressed by the whole document belonging to

said document group, A recording medium which recorded a program for making a computer perform a process of displaying an extracted this significant sentence on said screen as some documents belonging to said document group on a computer so that reading was possible.

[Translation done.]

*** NOTICES ***

JP0 and INPIT are not responsible for any damages caused by the use of this translation.

- 1.This document has been translated by computer. So the translation may not reflect the original precisely.
- 2.**** shows the word which can not be translated.
- 3.In the drawings, any words are not translated.

DETAILED DESCRIPTION

[Detailed Description of the Invention]**[0001]**

[Field of the Invention]This invention relates to the technology which displays at least some documents belonging to this document group on a screen so that an inspection is possible in detail about a document display system about the document group which comprised two or more documents which have a predetermined relation.

[0002]

[Description of the Prior Art]It is performed daily nowadays when communication technology progressed that two or more persons transmit information mutually using a character, a picture, etc. on networks, such as the Internet and personal computer communications. Such signal transduction like signal transduction other than on a network, It is materialized when those (receiver) who receive the information disseminated by the person (addresser) and this information addresser who disseminate information spontaneously towards others, and those who reply an information addresser the reply to the received information, an opinion, etc. (reply person) exist.

[0003]On the other hand, the signal transduction on a network has the feature that the exchange with an addresser, a receiver, or a reply person is performed via the server which provides service of transmission and reception of data, an inspection of received data, etc. That is, an addresser thru/or a reply person send to a server the information which it is going to disseminate thru/or reply via a network. Transmit the received information to a receiver via a network, or via a network, a receiver displays the received information so that an inspection is possible, or the server which received information saves it for the inspection by transmission to said receiver, or a receiver. The thing of those (an above-mentioned addresser and receiver, a reply person) who connect with a server and receive the above-mentioned service is hereafter called client.

[0004]The technique of performing signal transduction by sending each other's document between an addresser and a specific receiver as the technique of the signal transduction on a network in recent years. (For example, E-mail of a push type, etc.) The technique of performing signal transduction to except among many persons by the writing of the document to the electronic bulletin board (BBS) provided in the server is used. Specifically, the addresser who is going to send the document of a certain theme to others writes the contents of the document (henceforth a contribution document) which it is going to send in an electronic bulletin board. The inspection of the contents of the contribution document written in the electronic bulletin board is attained via a network at many persons (unspecified people or specific group all the members). The receiver who received the contents of the contribution document by inspection can write the document (henceforth the reply document to contribution) which makes the contents the answer to this contribution document, an opinion, etc. in an electronic bulletin board. The inspection of the contents of the reply document to the written-in contribution to many persons is attained like a contribution document. The receiver who received the contents of the reply document by inspection can write the document (henceforth the reply document to a reply) which makes the contents the answer to this reply document, an opinion, etc. in an electronic bulletin board. The inspection of the reply document to the written-in reply to many persons is attained like a contribution document etc.

[0005]The contribution document and the reply document to contribution have a series of relation as documents with which the theme is common.

It has a predetermined relation.

For this reason, in the conventional electronic bulletin board, it considered that a series of documents which have letter-sent-to-get-a-reply-reply relations were the document groups of 1, and the document belonging to each document group was arranged to the tree structure (layered structure which branched). the time of even the reply document located in the bottom of the heap of the tree structure to which this contribution document belongs from the contribution document of 1 located in the top layer of a tree structure here being connected due to letter-sent-to-get-a-reply-reply 1 -- this -- a settlement of all the documents connected due to letter-sent-to-get-a-reply-reply 1 is hereafter called thread.

[0006]The series chart of the document which belongs to each document group based on the above-mentioned tree structure is created, this series chart is displayed so that an inspection is possible, and the client enabled it to grasp the letter-sent-to-get-a-reply-reply relation between documents easily in the conventional electronic bulletin board. A client enabled it to grasp the outline of a document by writing together the beginning portions of the title of each document, or the text of each document in the series chart displayed, before accessing the text of each document.

[0007]

[Problem to be solved by the invention]However, in the signal transduction on the conventional network. In order it is difficult to grasp correctly the summary of each document which has a predetermined relation at a glance and to grasp the summary of each document correctly, He had to access the text of all the documents in an order from the document located in the top layer, and had to understand the contents of the text, and there was a problem that time and a labor will be applied to grasp of the summary of an exact document. When there were many document numbers especially displayed at once to a client, the above-mentioned problem was more remarkable.

[0008]For example, when it was an electronic bulletin board in which a series chart containing many documents is displayed, it was difficult to grasp an exact summary of a document from a title of each document displayed into a series chart. It is because it is hard to say that it is expressing an exact summary of a document in order that various addressers thru/or reply persons may indicate a title of each document in their different, free style. In particular, in a reply document by a reply person, in such a case, he was not able to understand a summary of a reply document at all from a title so that what added a loess display like "Re:" to the beginning of a title of a contribution document may be automatically made into a title. A summary of a document has not been grasped when it is indicated that it is not related to a summary of a document at the beginning, even if it wrote together a beginning portion of the text of each document on a series chart.

[0009]When many mails are received at once in a cellular phone etc., a majority of some of each mails (for example, beginning portions of a title or the text) are displayed on a display the first stage, but. In order to grasp a summary in which it is difficult and many mails are exact for the same Reason as the above-mentioned electronic bulletin board, grasping an exact summary of e-mail from a title of mail displayed this first stage a receiver, After an initial display, operation for displaying a mail text further was performed about all the mails, and a mail text had to be read, scrolling a screen where a mail text was displayed.

[0010]Also in the case of an E-mail etc. which send each other's document between an addresser and a specific receiver, transmitting mail and reception mail have a letter-sent-to-get-a-reply-reply relation, and it has a series of relation as a document with which the theme is common. About a technique grasped correctly at a glance, it was not proposed at all in the former, without accessing a summary of a series of mails which have such a letter-sent-to-get-a-reply-reply relation at each mail text.

[0011]This invention solved this problem, about a series of documents which have letter-sent-to-get-a-reply-reply relations, enabled grasp of a summary of each document promptly and exactly, and took the following composition for the purpose of enabling it to select document information simply and efficiently.

[0012]

[The means for solving a technical problem, and its operation and effect] A document group extraction means for a document display system of this invention to be a document display system which displays on a screen at least some two or more documents which have a predetermined relation so that an inspection is possible, and to extract a mass of document group from said two or more documents, A significant sentence extraction means to extract a sentence which has predetermined correlation in this document as a significant sentence from each document belonging to said document group in consideration of contents expressed by the whole document belonging to said document group, Let it be a summary to have had a significant sentence displaying means which displays a significant sentence extracted by this significant sentence extraction means on said screen as some documents belonging to said document group.

[0013]According to the document display system of the above-mentioned invention, a significant sentence extraction means takes into consideration contents expressed by the whole document belonging to a mass of document group, A sentence which has predetermined correlation in this document is extracted from each document belonging to a document group as a significant sentence, and a significant sentence displaying means displays an extracted significant sentence on a screen as some documents. Therefore, it can grasp correctly only by seeing a significant sentence displayed in deployment of the theme in a document group. It is also preferred to carry out two or more documents which have the above-mentioned predetermined relation a document top in relation between a letter sent to get a reply and a reply.

[0014]It is good also as composition which extracts a significant sentence according to a kind which was provided with a kind discriminating means which distinguishes a kind of document group, and was distinguished by this kind discriminating means. Since a significant sentence according to a kind of document group will be displayed if it carries out like this, thematic deployment can be grasped more correctly.

[0015]In this case, it is also preferred to use at least one of asked type of a question which includes question expression for a kind of document group, two or more subject type including two or more subjects, and interactive modes containing conversational sentences. It is also desirable to provide a significant sentence extraction means with a correlation setting-out means to change a standard which extracts a sentence which has predetermined correlation in this document as a significant sentence according to a kind of document belonging to this document group from each document belonging to a document group.

[0016]It is also desirable to have a summary preparing means which creates a summary of a document group based on a significant sentence extracted by a significant sentence extraction means, to change to a significant sentence displaying means and to have a summary displaying means which displays on a screen a summary created by this summary preparing

means. If it carries out like this, the contents of subject for every document group can be grasped promptly and exactly.

[0017]It does not interfere as composition provided with a significant sentence displaying means which displays a significant sentence extracted by a significant sentence extraction means with such a summary displaying means on said screen as some documents belonging to a document group.

[0018]It is also preferred to have a sentence standardization means which rewrites a sentence of each of said document to a standard expression based on a predetermined standard. If a summary preparing means has composition provided with a means which rewrites expression of a significant sentence extracted by said significant sentence extraction means at least to a standard expression, Since a sentence of a level uniform irrespective of skill about a document preparation person's document preparation is displayed as a summary of a document or a document group, it becomes much more easy to grasp the contents of subject for every document group.

[0019]As the above-mentioned sentence standardization means, short-sentence-izing of a redundant expression, substitution to other independent words of an independent word, correction of direction for use of an attached word, substitution from a dialect to the standard language, replacement to a character defined beforehand, declared unification of a shake, etc. can be considered, for example.

[0020]It is also preferred to have an owner meaning sentence extraction means to extract a sentence except expression which cannot constitute a summary of a document out of a document based on a predetermined standard as an owner meaning sentence, and to make a significant sentence extraction means or a summary preparing means into a means to perform extraction of a significant sentence or creation of a summary using this owner meaning sentence. As expression extracted by the above-mentioned expression extraction means, a citation part of a former utterance, a signature of a greeting sentence and an addresser, etc. can be considered, for example. If it carries out like this, extraction of a significant sentence or creation of a summary can be performed with efficiently and sufficient accuracy.

[0021]It is also preferred to make a significant sentence displaying means or a summary displaying means into a means which carries out the list display of a significant sentence about a document group or the summary to the order of a time series. If it carries out like this, the newest theme and subject can be accessed easily.

[0022]It is good also as a means to relate with a significant sentence or a summary a maker of a document which belongs a significant sentence displaying means or a summary displaying means to a document group, and to display him. If it carries out like this, it will become easy to grasp change of utterance contents by each maker.

[0023]It has a title preparing means which creates a title of a document group based on

extraction by a significant sentence extraction means, and even if a title of each document group created by this title preparing means is related with a significant sentence or a summary and it displays it, it does not interfere. If it carries out like this, as compared with a case where titles of each document attached by maker are enumerated, it will become easy to grasp a kind of subject of each document group.

[0024]A receiving set of this invention is a receiving set which receives a document via a communication line, and comprises the following:

A document group extraction means to extract a mass of document group from said two or more received documents.

This extraction means which takes out the text from each document belonging to this document group at least about a document group of 1 extracted by this document group extraction means.

An extraction means to extract a sentence which has predetermined correlation in this document as a significant sentence from each document belonging to said document group in consideration of contents expressed by the whole text taken out by this this extraction means.

An output means which outputs a significant sentence about said document group of 1 extracted by this extraction means via a communication line.

[0025]According to the receiving set of the above-mentioned invention, this extraction means takes out the text from each document in which a document group extraction means extracts a mass of document group from two or more documents received via a communication line, and belongs to a document group of extracted 1 at least. An extraction means extracts a sentence which has predetermined correlation in this document from each document which belongs to a document group in consideration of contents expressed by the taken-out whole text as a significant sentence, and an output means outputs a significant sentence about a document group of extracted 1 via a communication line. Therefore, it becomes possible about deployment of the theme in a document group of 1 to obtain correctly data which can be grasped by obtaining an outputted significant sentence via a communication line. For example, if it displays or this data is printed, it can grasp correctly only by seeing a significant sentence displayed thru/or printed in deployment of the theme in a document group.

[0026]It is also possible to divide above-mentioned document group means forming, this extraction means, extraction means, and output means into two or more pieces of equipment, to establish them, and to realize a receiving set of the above-mentioned invention combining two or more [these] pieces of equipment.

[0027]It is also desirable to have a preparing means which creates a summary of a document based on a significant sentence extracted by an extraction means, to change a summary output means which outputs a summary created by this preparing means via a communication

line to an output means, and to have it. If it carries out like this, utility value of data it becomes possible about the contents of subject for every document group to obtain data which can be grasped promptly and exactly can be raised by obtaining an outputted summary via a communication line.

[0028]A document display method of this invention is a document display method which displays on a screen at least some two or more documents which have a predetermined relation so that an inspection is possible, Extract a mass of document group from said two or more documents, and contents expressed by the whole document belonging to said document group are taken into consideration, Let it be a summary to extract a sentence which has predetermined correlation in this document as a significant sentence from each document belonging to said document group, and to display a this extracted significant sentence on said screen as some documents belonging to said document group.

[0029>About a mass of document group, in consideration of contents expressed by the whole document belonging to this document group, a sentence which has predetermined correlation in this document is extracted from each document as a significant sentence, and, according to the document display method of the above-mentioned invention, an extracted significant sentence is displayed on a screen as some documents. Therefore, it can grasp correctly only by seeing a significant sentence displayed in deployment of the theme in a document group.

[0030]A recording medium which recorded a computer program of this invention, A process of being the recording medium which recorded a computer program for displaying on a screen at least some two or more documents which have a predetermined relation so that an inspection is possible, and extracting a mass of document group from said two or more documents, A process of extracting a sentence which has predetermined correlation in this document as a significant sentence from each document belonging to said document group in consideration of contents expressed by the whole document belonging to said document group, Let it be a summary to have recorded a program for making a computer perform a process of displaying an extracted this significant sentence on said screen as some documents belonging to said document group on a computer so that reading was possible.

[0031]When a recorded program is read by computer according to the recording medium of the above-mentioned invention, In consideration of contents expressed by the whole document belonging to a mass of document group, a sentence which has predetermined correlation in this document is extracted from each document as a significant sentence, and an extracted significant sentence is displayed on a screen as some documents. Therefore, it can grasp correctly only by seeing a significant sentence displayed in deployment of the theme in a document group.

[0032]It is also possible to realize a recording medium of the above-mentioned invention by dividing into two or more recording media each process with which the above-mentioned

recording medium is provided, recording it, and combining two or more [these] recording media.

[0033]As a recording medium, various recording media (ROM, PROM, EEPROM, a flash memory, etc.), such as a flexible disk, CD-ROM, DVD-ROM, and semiconductor memory, can be used. These programs are memorized from the first to a server placed on networks, such as the Internet, and downloading and using for a computer of a client is also possible.

[0034]A display of a significant sentence, a summary, etc. can also be displayed by not remaining in what is displayed on a display or a paper, but reading out a significant sentence using voice synthesis etc. Since it is more difficult than reading and judging a character to distinguish immediately what is important when hearing it from an ear and recognizing the contents, composition which extracts and reads out a significant sentence is very useful. If it reads out a significant sentence with a sound in checking a mail document with a cellular phone, it is not restrained by small display of a cellular phone and is useful also at this point.

[0035]

[Mode for carrying out the invention]Hereafter, an embodiment of the invention is described based on an embodiment.

(1) Composition of an embodiment : first, use and explain drawing 1 about composition of an embodiment. Drawing 1 is an explanatory view showing outline composition of the bulletin board display system KS which is the 1st embodiment of this invention. it is shown in drawing 1 -- as -- the large-scale network 100 like [in the bulletin board display system KS] the Internet -- the distributing server 300,310,320 ... and the database server 200 are connected. the distributing server 300 -- the personal computer 900,910,920 of a client ... is connected via a telephone line. The personal computer 800 of a janitor of an electronic bulletin board is connected to the server 200 in which a web page of an electronic bulletin board was stored via a telephone line.

[0036]If an access signal to a web page of an electronic bulletin board is sent out from the personal computer 900 to the distributing server 300, a web page of an electronic bulletin board in the distributing server 300 and the database server 200 will be connected via the network 100. Thereby, a web page of an electronic bulletin board is displayed on a display of the personal computer 900. Then, when a client wishes printing of a document to an electronic bulletin board, a document which it is going to publish is drawn up using the personal computer 900, and a drawn-up document is sent to an address of an electronic bulletin board. A drawn-up document reaches a web page of an electronic bulletin board in the database server 200 via the network 100 from the distributing server 300. A document which a client drew up is written in an electronic bulletin board by this, and an electronic bulletin board after writing is displayed on a display of the personal computer 900. the personal computer 910,920 of other clients -- as well as the above-mentioned personal computer 900 when ... is used, access

thru/or writing to an electronic bulletin board is possible.

[0037]Structure of the database server 200 is shown in drawing 2. The database server 200, As the network interface (NT-I/F) 210 which controls an exchange of data with the network 100, CPU220 which perform processing, ROM230 which memorize a processing program and fixed data, and a work area. It has hard disk 270 grade which has memorized **RAM240, the timer 250 which manages time, the database accumulating part 260 which stores various kinds of data mentioned later, a dictionary, etc. Actually, the database accumulating part 260 shall be treated as independent equipment on account of explanation here, although stored in memory storage, such as a hard disk.

[0038]The document of a large number currently written in the electronic bulletin board is accumulated in the database accumulating part 260. These documents are accumulated in the state where it was arranged based on the letter-sent-to-get-a-reply-reply relation mentioned above. Specifically, the series chart of the document as shown in drawing 3 is stored in the database accumulating part 260 as information about a contribution document, the reply document to this contribution, and the reply document to a reply.

[0039]As shown in drawing 3, the letter-sent-to-get-a-reply-reply relation about each contribution document A-Z and a reply document is expressed in the series chart by the tree structure. For example, as a reply to the contribution document A, there are the four reply documents A-1-4, and there is two reply document A-1-a-b as a reply document to the reply document A-1. The date, a message number, an implementor name, etc. with which each document was written in the electronic bulletin board are appended to the series chart near each contribution document A-Z and the reply document (not shown). The text of each contribution documents A-Z and a reply document is linked with a series chart, and is accumulated in the database accumulating part 260.

[0040]A two-dot chain line showed an example of a thread mentioned above to drawing 3. That is, reply document A-1-a located in the 3rd layer that is the bottom of the heap of a tree structure to which this contribution document A belongs from the contribution document A of 1 located in the 1st layer that is the top layer of a tree structure is connected due to letter-sent-to-get-a-reply-reply 1 via the reply document A-1 located in the 2nd layer. A settlement (getting it blocked the contribution document A, the reply document A-1, reply document A-1-a) of all the documents connected by such a letter-sent-to-get-a-reply-reply relation serves as the thread TH. As long as it is expressed with a series chart of drawing 3 to drawing 3, "The contribution document A, the reply document A-1, reply document A-1-a", "The contribution document A, the reply document A-1, reply document A-1-b", "the contribution document A, the reply document A-2", The seven threads TH "the contribution document A, the reply document A-3, reply document A-3-a", "the contribution document A and the reply document A-4", "the contribution document B, the reply document B-1 and reply document B-1-a", and

"the contribution document Z, the reply document Z-1 and reply document Z-1-a" exist.

[0041]In ROM230 of the database server 200, a program which described processing (henceforth series chart creation processing) which creates the above-mentioned series chart is stored. This program functions as a sorting means which classifies many documents for every thread as a document group, when CPU220 performs. It is shown in drawing 4 by making contents and procedure of this series chart creation processing into a series chart creation manipulation routine. This routine is processing which CPU220 performs, when a contribution document and a reply document are written in. Starting of this routine will perform processing which specifies a document first made into an address of a written-in document in a series chart (Step S100). Next, processing which adds some written-in documents (for example, title etc.) is performed, relating with a document specified as a lower layer of a specified document (Step S120). Next, a contribution document and a reply document in a series chart are re(Step S140) classified for every thread, a series chart after a classification is updated and memorized to the database accumulating part 260 (Step S160), and this routine is ended.

[0042]In ROM230 of the database server 200. A summary of a contribution document written in an electronic bulletin board other than the above-mentioned series chart creation processing and a reply document is created, A program the contents of processing (henceforth summary creation and display processing) which displays this created summary on a display of the personal computer 900 with a series chart were described to be is stored. It is shown in drawing 5 by making contents and procedure of this summary creation and display processing into summary creation and a display-processing routine. This routine is processing which CPU220 performs, when execution instruction of a purport that a summary of a document written in an electronic bulletin board is created is made.

[0043]Starting of this routine will perform ** thread specific processing (Step S200) first. Thread specific processing is processing which specifies the thread TH of 1 which becomes an object which creates a summary out of a series chart. Henceforth, ** effective text specific processing (Step S210), ** significant sentence extracting processing (Step S220), ** summary creation processing (Step S240), and ** title creation processing (Step S260) are performed for this thread TH of specified 1. Details of these processings are mentioned later. When processing of these steps S220-S260 is completed about all the threads TH, (Step S270) and ** indicative-data output process (Step S280) mentioned later are performed, and this routine is ended.

[0044]Hereafter, the contents of ** effective text specific processing, ** significant sentence extracting processing, ** summary creation processing, ** title creation processing, and the ** indicative-data output process are explained in order. Although segmentation processing after whole evaluation is included in ** significant sentence extracting processing and sentence

standardization processing is included in ** summary creation processing, it combines with ** significant sentence extracting processing and ** summary creation processing, and these processings are also explained.

[0045]** Effective text specific processing (Step S210 of drawing 5)

Effective text specific processing is processing which specifies the text (henceforth the effective text) which extracted expression which cannot constitute a summary of a document from the text of each document which belongs in this thread TH about the thread TH specified by thread specific processing, and excepted extracted expression. In the case of the thread TH enclosed with a two-dot chain line by drawing 3, the text of the contribution document A is specifically read, a thin expression of a relation is excepted from the text to specification of the theme of documents, such as a greeting sentence and a contributor's signature, and the text after deletion is made into the effective text. Also with the reply document A-1 or reply document A-1-a, the text is read and the same processing as the contribution document A is performed. However, in the reply document A-1 or reply document A-1-a, it deletes from the text also about a part which quoted the text of the contribution document A, and the text after deletion is made into the effective text. Such the effective text is memorized temporarily in a predetermined region of RAM240.

[0046]An example of each document (the contribution document A, the reply document A-1, and reply document A-1-a) belonging to the thread TH enclosed with a two-dot chain line by drawing 3 is shown in drawing 6. The contribution document A is a document which was drawn up by the first and written in an electronic bulletin board, and is a document which the reply document A-1 was drawn up by the second as a reply to the contribution document A, and was written in an electronic bulletin board. Reply document A-1-a is the document which was drawn up by shell as a reply to the reply document A-1, and was written in an electronic bulletin board. As shown in drawing 6, four sentences, six sentences, and five sentences are contained in the reply document A-1, the reply document A-1, and reply document A-1-a, respectively, but. A finite greeting sentence of "being always indebted" of the 1st sentence of the reply document A-1 is deleted, and the five remaining sentences are made the effective text by execution of the above-mentioned effective text specific processing. in order to make it easy to know next explanation and to obtain, a statement number (1-14) of a sentence made into the effective text is shown in a beginning of a sentence of each sentence in drawing 6.

[0047]** Significant sentence extracting processing (Step S220 of drawing 5)

Significant sentence extracting processing is processing which extracts a sentence which has predetermined correlation in each document as a significant sentence from the effective text of each document specified by effective text specific processing. In consideration of contents expressed with this example in significant sentence extracting processing by the full-effective text which belongs in the thread TH of 1, as a sentence which has predetermined correlation in

each document, Processing (henceforth segmentation processing after whole evaluation) which starts a sentence of a meaning nearest to contents expressed by the full-effective text from the text of each document is performed. A sentence started by execution of this segmentation processing after whole evaluation is extracted as a significant sentence.

[0048] Although various techniques could be considered, the following technique was adopted as extracting processing of a significant sentence in this embodiment. Contents and a procedure of significant sentence extracting processing are shown in drawing 7 as a significant sentence extracting processing routine. This routine is started after an end of effective text specific processing. Starting of this routine will perform processing which distinguishes first a kind of each thread TH which an electronic bulletin board has (Step S300). It is for changing the extraction technique of a significant sentence from each document contained in the thread TH according to a kind of thread TH. (the actual contents of the document in a thread).

[whether as a kind of this thread, it is developed by the theme which differed from the theme of the first contribution document with each document in a thread, for example, and] Whether it is a thing of character whether it is a thread of a Q&A system, and in which a letter sent to get a reply and a reply are frequently repeated like a chat in (form of a thread), and a thread (Character of a thread, a special feature), or the length of one document in a thread is in which tendency of a long sentence or a short sentence (tendency of a document in a thread) -- etc. -- it can think. Suppose that a kind of thread TH is distinguished to asked type of a question, two or more subject type, or an interactive mode of three in this example.

[0049] Distinction of a kind of thread TH can be performed in the following ways. Hereafter, the thread TH enclosed with a two-dot chain line by drawing 3 is explained as an example. First, when the effective text of the contribution document A which is the first document that provides the theme is taken out and "a question of -" and expression of "- is not known" are in the effective text, this thread TH is distinguished as it is asked type of a question. In not being applied to asked type of a question, the full-effective text (effective text of the contribution document A, the reply document A-1, and reply document A-1-a) in the thread TH is taken out, and it judges whether two or more themes are within the limits of [fixed] the full-effective text. When this thread TH is distinguished as it is two or more subject type when there are two or more themes, and two or more themes cannot be found, this thread TH is distinguished as it is an interactive mode. Although a judgment of existence of two or more themes can be performed by changing the full-effective text into vectorial representation using the TFIDF method, detailed explanation is omitted here.

[0050] In this way, after distinguishing a kind of thread TH, processing which extracts a significant sentence according to a kind of thread TH is performed (Step S320). It attaches in detail [this processing], and divides and explains to asked type of a question, two or more subject type, and an interactive mode hereafter.

[0051]The extracting processing technique of a significant sentence in case a kind of thread TH is asked type of a question is shown in a flow chart of drawing 8. In being asked type of a question, the first document that provides the theme extracts first a sentence including "a question of -", and expression of "- is not known" from the effective text of ** (an example of drawing 3 the contribution document A), and let this be a significant sentence of the first document (Step S400, S410). Next, the first document and other documents (an example of drawing 3 reply document A-1-a) drawn up by the person same in the same thread TH as the first document are looked for, and a sentence including expression of "it succeeded" or "it having solved", "understanding", etc. is extracted from the effective text of other documents. An extracted sentence turns into a significant sentence of the first document and other documents drawn up by the same person (Step S420, S430). Then, a sentence in which solution to a question is contained in sentence order including expression of "having succeeded" of the effective text of a document besides the above is taken out, and a word contained in this sentence is started as a word which shows solution. Next, the first document extracts a sentence in which a word which shows the above-mentioned solution is contained from the effective text of others' document (an example drawing 3's reply document A-1) drawn up by different person. An extracted sentence turns into a significant sentence of others' document (Steps S440-S460).

[0052]Thus, in this example, when the kind of thread TH is asked type of a question. The sentence which has predetermined correlation in this document from each document in the thread TH (in this example.) The standard which starts the sentence of the meaning nearest to the contents expressed by the full-effective text is changed according to the kind (the example of drawing 3 the contribution document A, the reply document A-1, reply document A-1-a) of document belonging to this document group.

[0053]The extracting processing technique of a significant sentence in case the kinds of thread TH are two or more subject type and an interactive mode is shown in the flow chart of drawing 9. In the case of a subject type and an interactive mode, two or more processings which take out first the effective text (henceforth the full-effective text) of the whole sentence document which belongs in the thread TH of 1 are performed (Step S500).

[0054]In this example, it considered that the thread TH which has the example shown in drawing 6 was an interactive mode, and the extracting processing of the significant sentence shown in drawing 9 is applied. In the case of the example shown in drawing 6, a total of 14 sentences of the effective text (sentence of the statement numbers 5-9 shown in drawing 6) of the effective text (sentence of the statement numbers 1-4 shown in drawing 6) of the contribution document A and the reply document A-1 and the effective text (sentence of the statement numbers 10-14 shown in drawing 6) of reply document A-1-a are taken out as the full-effective text.

[0055]In a two or more subject type case, in processing of Step S500, the text which belongs to a certain theme of 1 among the full-effective texts, and the text belonging to other themes are taken out independently, make each taken-out text the full-effective text, and processing not more than step S510 is performed.

[0056]Next, a word which is a noun is extracted from the full-effective text, and an extracted word performs processing which searches for a grade which inclines and occurs frequently in each sentence which constitutes the full-effective text (Step S510).

[0057]The grade in which a word inclines and occurs frequently can evaluate the number of times to which the word appears within the full-effective text by a value normalized with the number of sentences which constitute the full-effective text. This is known, for example as TFIDF. TFIDF is defined by the following formula. Each sentence in which db is a mass of target text data (here full-effective text), and d constitutes the full-effective text from the following formulas, and t are taken as a word contained in this text.

[0058]

$$TFIDF(d, t) = TF(d, t) \times IDF(t) \quad -- (1)$$

however, : -- TF (d, t) depends the number of times and ldf to which a word "t" appears in each sentence d on a following formula (2).

$$IDF(t) = \text{LOGe} \{DB(db)/f(t, db)\} \quad -- (2)$$

the number of sentences with which, as for the value 14 and f (t, db), a word "t" appears in the full-effective text here in the number of sentences with which DB (db) constitutes the full-effective text, therefore an example which showed drawing 6 -- it comes out.

[0059]In an example shown in drawing 6, drawing 10 shows a result of having counted the frequency of occurrence of a word which is the noun contained in this example. In this way, after asking for the frequency of occurrence of a word which is a noun, in this example, the frequency of occurrence calculated a TFIDF value about two or more words based on an upper type. In this way, a TFIDF value of each called-for word is shown in drawing 11.

[0060]Then, processing which adds together TFIDF of a word which is a noun which constitutes each sentence is performed (Step S520). A result of this total processing was shown in a column of a total value of drawing 11.

[0061]Next, size of a total value of each sentence is judged for every effective text of each document contained in the thread TH, and processing which extracts a sentence which is two with the largest total value is performed (Step S530). As with an asterisk shows to drawing 11, in the case of an example of drawing 6, a sentence of the statement numbers 7 and 8 from the reply document A-1 and a sentence of the statement numbers 11 and 14 from reply document A-1-a are extracted for a sentence of the statement numbers 1 and 2 from the contribution document A, respectively.

[0062]Processing to the above steps S510-S530 is equivalent to the segmentation processing

after whole evaluation mentioned above. According to this segmentation processing after whole evaluation, if the frequency of occurrence within each effective text is not high, the TFIDF value of a word does not become high. On the other hand, in the case of a word (in for example, the in "things" and the "case") which appears in the full-effective text uniformly, since $f(t, db)$ becomes a big value, a TFIDF value becomes with a value small after all. That is, it becomes a big value as the sentence in which the word which appears by high frequency in each sentence which constitutes each effective text exists, and moreover, it turns into a value with a high total value, so that a TFIDF value has many such a words.

[0063]As mentioned above, two significant sentences are extracted from each document contained in the thread TH. Then, the extracted significant sentence is memorized to the predetermined region of RAM240, this (Step S340 of drawing 7) routine is ended, and it shifts to summary creation processing (Step S240 of drawing 5).

[0064]** Summary creation processing (Step S240 of drawing 5)

They are summary creation processing, next the processing which creates a summary of the thread TH based on a significant sentence extracted by significant sentence extracting processing. In an example of drawing 3, a summary of the thread TH which consists of the contribution document A, the reply document A-1, and reply document A-1-a is created. Creation of this summary is created by specifically combining a significant sentence extracted from each document in a thread based on a predetermined algorithm. For example, the thread TH enclosed with a two-dot chain line is asked type of a question, the contribution document A includes a question in drawing 3, and the reply document A-1 includes a proposal of solution in it, What is necessary is to enlarge weight which combines a significant sentence extracted from the contribution document A, and a significant sentence extracted from reply document A-1-a, and just to create a summary, in including a result that reply document A-1-a was solved. It is because it is thought that a sentence "the point in question -- was solved" makes a visitor produce interest generally.

[0065]In this example, the summary is created by performing processing (henceforth sentence standardization processing) which rewrites the significant sentence extracted by the significant sentence by significant sentence extracting processing to a standard expression. Specifically in sentence standardization processing, short-sentence-izing of a redundant expression, the substitution to other independent words of an independent word, correction of the direction for use of an attached word, the substitution from a dialect to the standard language, the replacement to the character defined beforehand, declared unification of a shake, etc. are processed. If it divides roughly into processing of standardization, standardization of a character, declared unification of a shake, unification of an independent word, etc. can be considered. Although explanation is omitted about these processings, it is as follows when rough illustration is mentioned.

[0066]If standardization of a character is illustrated, replacement of (A) parenthesis:"" and "" will be performed, (C) general [replacing quotation-marks: ""and"" etc.] sign : (B) About various signs (for example, ":", "?!", etc.). (D) katakana, such as replacing half width/full width, and an alphanumeric character : Replacement of full width/half width, and a capital letter/small letter is performed, (E) period and punctuation-marks: -- a period and punctuation marks -- "-- " -- "." -- *****, such as transposing *****: "Queen = Elizabeth" of the (F) name to "Queen Elizabeth", such as unifying.

[0067]The shake of the notation means the ambiguity of the notation in Japanese, and allowable width. For example, the shake of the example of shake:of ** macron, Windows, Windows, and ** declensional kana ending : An example, the example of shake:of the shake:angel of the example of shake:of the kana notation of sales, sales, the example of shake:of ** contracted sound notation, a wizard, a wizard, and ** compound, sales, sales, and ** loanword transcription, an angel, and ** iteration statement character, and a right right grand one -- **** can be mentioned fairly.

[0068]furthermore -- as unification processing of an independent word -- (b) modifier: -- ****, such as (**) Verb: telling about, connecting, and notifying a (**) noun:personal computer -- it is uncanny, the first and high -- a personal computer, PC, etc., can be illustrated.

[0069]Criterion data and a rule for performing such processing are stored in a dictionary in the hard disk 270 of the database server 200. If replacement to a character defined beforehand is taken for an example, the word "PC" will be transposed to the word a "personal computer." Thereby, a document in the thread TH of 1 and a term of a significant sentence can be unified.

[0070]** Title creation processing (Step S260 of drawing 5)

Title creation processing is processing which creates a title of a thread based on a significant sentence extracted by significant sentence extracting processing. In an example of drawing 3, a title of the thread TH which consists of the contribution document A, the reply document A-1, and reply document A-1-a is created. A title is created by analyzing grammar information on a significant sentence, specifically starting a required word, and complementing an attached word.

[0071]** An indicative-data output process (Step S280 of drawing 5)

An indicative-data output process is processing which outputs a summary created [above-mentioned], a title, and a series chart as an indicative data. This processing is performed when it is judged that a summary and a title were created about all the threads TH of an electronic bulletin board in processing of Step S270 of drawing 5.

[0072]An example of a data image outputted in an indicative-data output process is shown in drawing 12. A data image based on a series chart shown in drawing 3 is expressed with drawing 12. This data image is contents displayed on a preview screen of a display of the personal computer 900 the first stage, when an indicative data reaches the personal computer

900 of a client via the network 100 from the database server 200.

[0073]As shown in drawing 12, the summary and the title are displayed on the preview screen for every thread. The thread name of the thread TH which the thread name is attached based on the reply document located in the bottom of the heap of a series chart, for example, was enclosed with the two-dot chain line by drawing 3 is set to "A-1-a" which is a reply document name of the bottom of the heap. In this example, since they are collectively displayed by a contribution day (day when the contribution document was written in), an updated date (the latest day when the reply document was written in), and the hierarchy number (the number of the reply documents in a thread), these indication can be given the reference works about selection of subject.

[0074]The list display of the title about each thread and the summary is carried out to the order of a time series. In this example, the updated date shows sequentially from a new thread. If it says by drawing 12, a thread name will be the thread by which the thing of "B-1-a" was updated most these days. Thereby, the newest theme and subject can be accessed easily.

[0075]In a preview screen shown in drawing 12, it is good also as displaying a name of a maker of a contribution document belonging to each thread besides being a title, a summary, etc., or a reply document. If it carries out like this, it will become easy to grasp change of utterance contents by each maker.

[0076]The client can choose a thread of subject which is interested with reference to a title or a summary in a preview screen. In this way, selection of a thread will display all the document information which belongs in a selected thread on a list display screen with a letter-sent-to-get-a-reply-reply relation of a document. **A thread name shows drawing 13 contents displayed on a list display screen, when the thread TH of "A-1-a" is chosen. With this list display screen, the contents of interested subject can be checked in detail.**

[0077]The bulletin board display system KS of this example described above, Contents expressed by the whole document belonging to the thread TH are taken into consideration about the thread TH which comprised a contribution document which has letter-sent-to-get-a-reply-reply relations, a reply document to contribution, and a reply document to a reply, A sentence of a meaning nearest to the contents of each document to the whole document belonging to the thread TH is extracted as a significant sentence, a summary of the thread TH is created based on an extracted significant sentence, and a created summary is displayed on a screen. Therefore, it becomes possible to grasp the contents of subject for every thread TH promptly and exactly, and subject can be selected promptly and easily.

[0078]The bulletin board display system KS of this example distinguishes a kind of thread TH, extracts a significant sentence according to a kind of distinguished thread TH, and displays this significant sentence. Therefore, thematic deployment and the contents of subject can be grasped more correctly.

[0079]The bulletin board display system KS of this example applies a sentence standardization means to an extracted significant sentence, rewrites a significant sentence to a standard expression, and creates a summary. Since a sentence of a level uniform irrespective of skill about a document preparation person's document preparation will be displayed as a summary of a thread if it carries out like this, it becomes much more easy to grasp the contents of subject for every document group. For example, when a significant sentence which has a redundant expression is short-sentence-ized, a sentence which the contents tend to grasp in a limited viewing area can be displayed.

[0080]The bulletin board display system KS of this example performs extraction of a significant sentence, or creation of a summary based on the effective text which excepted expression which cannot constitute a summary of a document to a document belonging to each thread TH. Therefore, extraction of a significant sentence or creation of a summary can be performed with efficiently and sufficient accuracy.

[0081]The bulletin board display system KS of this example creates a title of the thread TH using an extraction result of a significant sentence, and displays it with a summary of the thread TH. Therefore, as compared with a case where titles of each document attached by maker are enumerated, it becomes easy to grasp a kind of subject of each thread TH.

[0082](2) The 2nd embodiment : there is an E-mail which sends each other's document between an addresser and a specific receiver in addition to a write-in document to an electronic bulletin board in a document which has letter-sent-to-get-a-reply-reply relations. A significant sentence is extracted to a settlement of such an E-mail, and an embodiment which creates a summary is described as an **e-mail display system MS**. Drawing 14 is an explanatory view showing outline composition of e-mail display system MS which is the 2nd embodiment of this invention. As shown in drawing 14, although e-mail display system MS is provided with the bulletin board display system KS and almost common composition, it differs from the bulletin board display system KS in that it does not have the database server 200. Double figures are expressed with drawing 14 about this common composition using the same number as drawing 1 on a mark.

[0083]The program which described the processing which arranges a settlement (henceforth an e-mail bunch) of the mail transmitted and received between the specific addresser and the specific receiver in order of reception thru/or a transmission date, and creates a time sequence diagram table is stored in ROM of the distributing servers 30 and 31. Execution of this program will create a time sequence diagram table as shown in drawing 15. As shown in drawing 15, Mr. Y who received the mails X1 transmitted to Mr. Y from X persons transmitted the mail Y1 to X persons, and X person who received this mail Y1 has transmitted the mail X2 to Mr. Y. In drawing 15, transmission and reception of mail between such X person and Mr. Y continue 5 times. This has chained the mail exchanged between X persons and Mr. Y in order of a time

series like "the e-mail X1 -> mail Y1 -> mail X2 -> mail Y1 -> mail X3 -> mail Y3 -> mail X4 -> mail Y4 -> mail X5 -> mail Y5." A total of ten copies of mails chained in this way becomes the e-mail bunch MB. The text of each mail in a time sequence diagram table is linked with a time sequence diagram table, and is accumulated in the hard disk in the distributing server 30 and 31.

[0084]A program the same contents of summary creation and display processing as the 1st above-mentioned embodiment were described to be is stored in ROM of the distributing servers 30 and 31. By executing this program with the command of CPU of the distributing servers 30 and 31, ** thread specific processing, effective text specific processing, ** significant sentence extracting processing, ** summary creation processing, ** title creation processing, ** indicative-data output process, and same processing are performed like the thread TH in the 1st embodiment to an e-mail bunch. As a result, a sentence of a meaning nearest to the contents of the whole mail is extracted from each mail which belongs to an e-mail bunch in consideration of contents expressed by the whole mail belonging to an e-mail bunch as a significant sentence, and a summary of an e-mail bunch is created based on an extracted significant sentence. A summary of such an e-mail bunch is displayed on X persons' preview screen of both sides the personal computer 90 and Mr. Y's personal computer 91. Therefore, a client becomes possible [grasping promptly and exactly the contents of subject for every e-mail bunch], and can select e-mail promptly and easily.

[0085]In the above-mentioned embodiment, although a display of the text of e-mail was taken up as an example, the technique of a such display is not limited to a display of a mail text, and can be applied also to data which comprised two or more sentences. For example, a significant sentence can be further extracted out of an abstract sentence for an abstract sentence of a paper database, and it can realize in a mode of displaying this with a title of a paper. Or real original form voice of extracting and displaying a significant sentence from a description of an electronic chart and reports, such as a newspaper distributed electronically, also exists. It is preferred for an extracted significant sentence it not only to to display on a monitor, but to replace with a display on a monitor or to read out with a sound with a display. From the first, printing with a printer etc. does not interfere, either.

[0086]As mentioned above, although an embodiment of the invention was described, As for the ability to carry out with a form which becomes various within limits which do not deviate from a summary of this invention, it is needless to say that this invention is not limited to such an embodiment at all, and displays a significant sentence in a different mode (for example, different colors) from others, for example etc.

[0087]Although it presupposed that the database server 200 performs series chart creation processing (drawing 4), and summary creation and display processing (drawing 5) in the 1st embodiment, It is good also as composition which stores in equipment other than database

server 200 a program which described the contents of series chart creation processing, or summary creation and display processing, and equipment other than database server 200 performs. For example, composition etc. which the distributing server 300,310,320 and the personal computer 900,910,920 of a client perform can be considered.

[0088]Although it presupposed that the distributing servers 30 and 31 perform series chart creation processing, and summary creation and display processing in the 2nd embodiment, It is good also as composition which stores in the distributing server 30 and equipment other than 31 a program which described the contents of series chart creation processing, or summary creation and display processing, and the distributing server 30 and equipment other than 31 perform. For example, composition etc. which the personal computers 90 and 91 which read the above-mentioned program perform can be considered.

[0089]In the above-mentioned embodiment, in significant sentence extracting processing, a write-in document to an electronic bulletin board, Although the number of average sentences is about five and Kazufumi determined the number of sentences to extract as two in consideration of being in a comparatively short tendency (processing of Step S530 of drawing 9), the number of sentences to extract, According to the length etc. of each document (if it says by drawing 3, they will be the contribution document A, the reply document A-1, and reply document A-1-a) contained in the thread TH, it can set arbitrarily.

[0090]In the above-mentioned embodiment, after extracting a significant sentence, sentence standardization processing was performed to this significant sentence, but before extracting a significant sentence, it is good also as composition which performs sentence standardization processing to the effective texts (text of a contribution document or a reply document, etc.).

[0091]Although summary creation processing (Step S240 of drawing 5) and title creation processing (Step S260 of drawing 5) were performed in the above-mentioned embodiment after an end of significant sentence extracting processing (Step S220 of drawing 5), It is good also as composition which outputs a significant sentence of each document extracted by significant sentence extracting processing by an indicative-data output process (Step S280 of drawing 5) so that a display is possible, without performing this summary creation processing and title creation processing. A sentence of a meaning nearest to the contents of each document to the whole document which belongs to the thread TH also by this composition in consideration of contents expressed by the whole document belonging to a thread TH mail bunch is extracted as a significant sentence, and an extracted significant sentence is displayed on a screen. Therefore, it can grasp correctly only by seeing a significant sentence displayed in deployment of the theme in each thread TH or each e-mail bunch.

[0092]When a significant sentence is extracted from each mail which belongs to an e-mail bunch especially and this significant sentence is arranged to a time series, By displaying on a monitor a significant sentence arranged in this way, or printing out, it becomes possible to

grasp the past exchange at a glance, and can utilize effectively in a business scene etc. specific contents (for example, delivery date when both sides have agreed) in the past exchange -- investigation -- when like, it is because it is not necessary to say that each mail belonging to an e-mail bunch is displayed on a monitor one by one, or it prints out, and specific contents are looked for from an e-mail bunch.

[0093]Although some embodiments and modifications of this invention were explained above, this invention is not limited to these embodiments at all, and can be carried out in various modes within limits which do not deviate from a summary of this invention. For example, a relation of two or more documents is not restricted to a relation of a letter-sent-to-get-a-reply-reply, and even if they are two or more documents, a comment, a diary, a report that the same maker drew up, for example, it does not interfere. A portion which performs a document display of this invention may be placed on a server which is accumulating two or more documents, and may be set to a client side which peruses two or more documents. Or it is also possible to place in the middle of these client/server architectures as a server for exclusive use. A program for constituting these systems, It is good also as what builds at least some document display systems which it could be dealt with with a form of media, such as CD-ROM, and what places a program on a server and is used downloaded this, and were mentioned above on their computer.

[Translation done.]

*** NOTICES ***

JP0 and INPIT are not responsible for any damages caused by the use of this translation.

- 1.This document has been translated by computer. So the translation may not reflect the original precisely.
- 2.**** shows the word which can not be translated.
- 3.In the drawings, any words are not translated.

TECHNICAL FIELD

[Field of the Invention]This invention relates to the technology which displays at least some documents belonging to this document group on a screen so that an inspection is possible in detail about a document display system about the document group which comprised two or more documents which have a predetermined relation.

[Translation done.]

* NOTICES *

JP0 and INPIT are not responsible for any damages caused by the use of this translation.

- 1.This document has been translated by computer. So the translation may not reflect the original precisely.
- 2.**** shows the word which can not be translated.
- 3.In the drawings, any words are not translated.

PRIOR ART

[Description of the Prior Art]It is performed daily nowadays when communication technology progressed that two or more persons transmit information mutually using a character, a picture, etc. on networks, such as the Internet and personal computer communications. Such signal transduction like signal transduction other than on a network, It is materialized when those (receiver) who receive the information disseminated by the person (addresser) and this information addresser who disseminate information spontaneously towards others, and those who reply an information addresser the reply to the received information, an opinion, etc. (reply person) exist.

[0003]On the other hand, the signal transduction on a network has the feature that the exchange with an addresser, a receiver, or a reply person is performed via the server which provides service of transmission and reception of data, an inspection of received data, etc. That is, an addresser thru/or a reply person send to a server the information which it is going to disseminate thru/or reply via a network. Transmit the received information to a receiver via a network, or via a network, a receiver displays the received information so that an inspection is possible, or the server which received information saves it for the inspection by transmission to said receiver, or a receiver. The thing of those (an above-mentioned addresser and receiver, a reply person) who connect with a server and receive the above-mentioned service is hereafter called client.

[0004]The technique of performing signal transduction by sending each other's document between an addresser and a specific receiver as the technique of the signal transduction on a network in recent years. (For example, E-mail of a push type, etc.) The technique of performing signal transduction to except among many persons by the writing of the document to the electronic bulletin board (BBS) provided in the server is used. Specifically, the addresser who is going to send the document of a certain theme to others writes the contents of the document (henceforth a contribution document) which it is going to send in an electronic

bulletin board. The inspection of the contents of the contribution document written in the electronic bulletin board is attained via a network at many persons (unspecified people or specific group all the members). The receiver who received the contents of the contribution document by inspection can write the document (henceforth the reply document to contribution) which makes the contents the answer to this contribution document, an opinion, etc. in an electronic bulletin board. The inspection of the contents of the reply document to the written-in contribution to many persons is attained like a contribution document. The receiver who received the contents of the reply document by inspection can write the document (henceforth the reply document to a reply) which makes the contents the answer to this reply document, an opinion, etc. in an electronic bulletin board. The inspection of the reply document to the written-in reply to many persons is attained like a contribution document etc.

[0005]The contribution document and the reply document to contribution have a series of relation as documents with which the theme is common.

It has a predetermined relation.

For this reason, in the conventional electronic bulletin board, it considered that a series of documents which have letter-sent-to-get-a-reply-reply relations were the document groups of 1, and the document belonging to each document group was arranged to the tree structure (layered structure which branched). the time of even the reply document located in the bottom of the heap of the tree structure to which this contribution document belongs from the contribution document of 1 located in the top layer of a tree structure here being connected due to letter-sent-to-get-a-reply-reply 1 -- this -- a settlement of all the documents connected due to letter-sent-to-get-a-reply-reply 1 is hereafter called thread.

[0006]The series chart of the document which belongs to each document group based on the above-mentioned tree structure is created, this series chart is displayed so that an inspection is possible, and the client enabled it to grasp the letter-sent-to-get-a-reply-reply relation between documents easily in the conventional electronic bulletin board. A client enabled it to grasp the outline of a document by writing together the beginning portions of the title of each document, or the text of each document in the series chart displayed, before accessing the text of each document.

[Translation done.]

*** NOTICES ***

JP0 and INPIT are not responsible for any damages caused by the use of this translation.

- 1.This document has been translated by computer. So the translation may not reflect the original precisely.
- 2.**** shows the word which can not be translated.
- 3.In the drawings, any words are not translated.

EFFECT OF THE INVENTION

[The means for solving a technical problem, and its operation and effect] A document group extraction means for the document display system of this invention to be a document display system which displays on a screen at least some two or more documents which have a predetermined relation so that an inspection is possible, and to extract a mass of document group from said two or more documents, A significant sentence extraction means to extract the sentence which has predetermined correlation in this document as a significant sentence from each document belonging to said document group in consideration of the contents expressed by the whole document belonging to said document group, Let it be a summary to have had the significant sentence displaying means which displays the significant sentence extracted by this significant sentence extraction means on said screen as some documents belonging to said document group.

[0013]According to the document display system of the above-mentioned invention, a significant sentence extraction means takes into consideration the contents expressed by the whole document belonging to a mass of document group, The sentence which has predetermined correlation in this document is extracted from each document belonging to a document group as a significant sentence, and a significant sentence displaying means displays the extracted significant sentence on a screen as some documents. Therefore, it can grasp correctly only by seeing the significant sentence displayed in deployment of the theme in a document group. It is also preferred to carry out two or more documents which have the above-mentioned predetermined relation the document top in the relation between a letter sent to get a reply and a reply.

[0014]It is good also as composition which extracts a significant sentence according to a kind which was provided with a kind discriminating means which distinguishes a kind of document group, and was distinguished by this kind discriminating means. Since a significant sentence according to a kind of document group will be displayed if it carries out like this, thematic

deployment can be grasped more correctly.

[0015]In this case, it is also preferred to use at least one of asked type of a question which includes question expression for a kind of document group, two or more subject type including two or more subjects, and interactive modes containing conversational sentences. It is also desirable to provide a significant sentence extraction means with a correlation setting-out means to change a standard which extracts a sentence which has predetermined correlation in this document as a significant sentence according to a kind of document belonging to this document group from each document belonging to a document group.

[0016]It is also desirable to have a summary preparing means which creates a summary of a document group based on a significant sentence extracted by a significant sentence extraction means, to change to a significant sentence displaying means and to have a summary displaying means which displays on a screen a summary created by this summary preparing means. If it carries out like this, the contents of subject for every document group can be grasped promptly and exactly.

[0017]It does not interfere as composition provided with a significant sentence displaying means which displays a significant sentence extracted by a significant sentence extraction means with such a summary displaying means on said screen as some documents belonging to a document group.

[0018]It is also preferred to have a sentence standardization means which rewrites a sentence of each of said document to a standard expression based on a predetermined standard. If a summary preparing means has composition provided with a means which rewrites expression of a significant sentence extracted by said significant sentence extraction means at least to a standard expression, Since a sentence of a level uniform irrespective of skill about a document preparation person's document preparation is displayed as a summary of a document or a document group, it becomes much more easy to grasp the contents of subject for every document group.

[0019]As the above-mentioned sentence standardization means, short-sentence-izing of a redundant expression, substitution to other independent words of an independent word, correction of direction for use of an attached word, substitution from a dialect to the standard language, replacement to a character defined beforehand, declared unification of a shake, etc. can be considered, for example.

[0020]It is also preferred to have an owner meaning sentence extraction means to extract a sentence except expression which cannot constitute a summary of a document out of a document based on a predetermined standard as an owner meaning sentence, and to make a significant sentence extraction means or a summary preparing means into a means to perform extraction of a significant sentence or creation of a summary using this owner meaning sentence. As expression extracted by the above-mentioned expression extraction means, a

citation part of a former utterance, a signature of a greeting sentence and an addresser, etc. can be considered, for example. If it carries out like this, extraction of a significant sentence or creation of a summary can be performed with efficiently and sufficient accuracy.

[0021]It is also preferred to make a significant sentence displaying means or a summary displaying means into a means which carries out the list display of a significant sentence about a document group or the summary to the order of a time series. If it carries out like this, the newest theme and subject can be accessed easily.

[0022]It is good also as a means to relate with a significant sentence or a summary a maker of a document which belongs a significant sentence displaying means or a summary displaying means to a document group, and to display him. If it carries out like this, it will become easy to grasp change of utterance contents by each maker.

[0023]It has a title preparing means which creates a title of a document group based on extraction by a significant sentence extraction means, and even if a title of each document group created by this title preparing means is related with a significant sentence or a summary and it displays it, it does not interfere. If it carries out like this, as compared with a case where titles of each document attached by maker are enumerated, it will become easy to grasp a kind of subject of each document group.

[0024]A receiving set of this invention is a receiving set which receives a document via a communication line, and comprises the following:

A document group extraction means to extract a mass of document group from said two or more received documents.

This extraction means which takes out the text from each document belonging to this document group at least about a document group of 1 extracted by this document group extraction means.

An extraction means to extract a sentence which has predetermined correlation in this document as a significant sentence from each document belonging to said document group in consideration of contents expressed by the whole text taken out by this this extraction means.

An output means which outputs a significant sentence about said document group of 1 extracted by this extraction means via a communication line.

[0025]According to the receiving set of the above-mentioned invention, this extraction means takes out the text from each document in which a document group extraction means extracts a mass of document group from two or more documents received via a communication line, and belongs to a document group of extracted 1 at least. An extraction means extracts a sentence which has predetermined correlation in this document from each document which belongs to a document group in consideration of contents expressed by the taken-out whole text as a significant sentence, and an output means outputs a significant sentence about a document

group of extracted 1 via a communication line. Therefore, it becomes possible about deployment of the theme in a document group of 1 to obtain correctly data which can be grasped by obtaining an outputted significant sentence via a communication line. For example, if it displays or this data is printed, it can grasp correctly only by seeing a significant sentence displayed thru/or printed in deployment of the theme in a document group.

[0026]It is also possible to divide above-mentioned document group means forming, this extraction means, extraction means, and output means into two or more pieces of equipment, to establish them, and to realize a receiving set of the above-mentioned invention combining two or more [these] pieces of equipment.

[0027]It is also desirable to have a preparing means which creates a summary of a document based on a significant sentence extracted by an extraction means, to change a summary output means which outputs a summary created by this preparing means via a communication line to an output means, and to have it. If it carries out like this, utility value of data it becomes possible about the contents of subject for every document group to obtain data which can be grasped promptly and exactly can be raised by obtaining an outputted summary via a communication line.

[0028]A document display method of this invention is a document display method which displays on a screen at least some two or more documents which have a predetermined relation so that an inspection is possible, Extract a mass of document group from said two or more documents, and contents expressed by the whole document belonging to said document group are taken into consideration, Let it be a summary to extract a sentence which has predetermined correlation in this document as a significant sentence from each document belonging to said document group, and to display a this extracted significant sentence on said screen as some documents belonging to said document group.

[0029>About a mass of document group, in consideration of contents expressed by the whole document belonging to this document group, a sentence which has predetermined correlation in this document is extracted from each document as a significant sentence, and, according to the document display method of the above-mentioned invention, an extracted significant sentence is displayed on a screen as some documents. Therefore, it can grasp correctly only by seeing a significant sentence displayed in deployment of the theme in a document group.

[0030]A recording medium which recorded a computer program of this invention, A process of being the recording medium which recorded a computer program for displaying on a screen at least some two or more documents which have a predetermined relation so that an inspection is possible, and extracting a mass of document group from said two or more documents, A process of extracting a sentence which has predetermined correlation in this document as a significant sentence from each document belonging to said document group in consideration of contents expressed by the whole document belonging to said document group, Let it be a

summary to have recorded a program for making a computer perform a process of displaying an extracted this significant sentence on said screen as some documents belonging to said document group on a computer so that reading was possible.

[0031]When a recorded program is read by computer according to the recording medium of the above-mentioned invention, In consideration of contents expressed by the whole document belonging to a mass of document group, a sentence which has predetermined correlation in this document is extracted from each document as a significant sentence, and an extracted significant sentence is displayed on a screen as some documents. Therefore, it can grasp correctly only by seeing a significant sentence displayed in deployment of the theme in a document group.

[0032]It is also possible to realize a recording medium of the above-mentioned invention by dividing into two or more recording media each process with which the above-mentioned recording medium is provided, recording it, and combining two or more [these] recording media.

[0033]As a recording medium, various recording media (ROM, PROM, EEPROM, a flash memory, etc.), such as a flexible disk, CD-ROM, DVD-ROM, and semiconductor memory, can be used. These programs are memorized from the first to a server placed on networks, such as the Internet, and downloading and using for a computer of a client is also possible.

[0034]A display of a significant sentence, a summary, etc. can also be displayed by not remaining in what is displayed on a display or a paper, but reading out a significant sentence using voice synthesis etc. Since it is more difficult than reading and judging a character to distinguish immediately what is important when hearing it from an ear and recognizing the contents, composition which extracts and reads out a significant sentence is very useful. If it reads out a significant sentence with a sound in checking a mail document with a cellular phone, it is not restrained by small display of a cellular phone and is useful also at this point.

[0035]

[Mode for carrying out the invention]Hereafter, an embodiment of the invention is described based on an embodiment.

(1) Composition of an embodiment : first, use and explain drawing 1 about composition of an embodiment. Drawing 1 is an explanatory view showing outline composition of the bulletin board display system KS which is the 1st embodiment of this invention. it is shown in drawing 1 -- as -- the large-scale network 100 like [in the bulletin board display system KS] the Internet -- the distributing server 300,310,320 ... and the database server 200 are connected. the distributing server 300 -- the personal computer 900,910,920 of a client ... is connected via a telephone line. The personal computer 800 of a janitor of an electronic bulletin board is connected to the server 200 in which a web page of an electronic bulletin board was stored via a telephone line.

[0036]If the access signal to the web page of an electronic bulletin board is sent out from the personal computer 900 to the distributing server 300, the web page of the electronic bulletin board in the distributing server 300 and the database server 200 will be connected via the network 100. Thereby, the web page of an electronic bulletin board is displayed on the display of the personal computer 900. Then, when a client wishes printing of the document to an electronic bulletin board, the document which it is going to publish is drawn up using the personal computer 900, and the drawn-up document is sent to the address of an electronic bulletin board. The drawn-up document reaches the web page of the electronic bulletin board in the database server 200 via the network 100 from the distributing server 300. The document which the client drew up is written in an electronic bulletin board by this, and the electronic bulletin board after writing is displayed on the display of the personal computer 900. the personal computer 910,920 of other clients -- as well as the above-mentioned personal computer 900 when ... is used, access thru/or the writing to an electronic bulletin board is possible.

[0037]The structure of the database server 200 is shown in drawing 2. The database server 200, As the network interface (NT-I/F) 210 which controls an exchange of data with the network 100, CPU220 which perform processing, ROM230 which memorize a processing program and fixed data, and a work area. It has the hard disk 270 grade which has memorized **RAM240, the timer 250 which manages time, the database accumulating part 260 which stores various kinds of data mentioned later, the dictionary, etc. Actually, the database accumulating part 260 shall be treated as independent equipment on account of explanation here, although stored in memory storage, such as a hard disk.

[0038]The document of a large number currently written in the electronic bulletin board is accumulated in the database accumulating part 260. These documents are accumulated in the state where it was arranged based on the letter-sent-to-get-a-reply-reply relation mentioned above. Specifically, the series chart of the document as shown in drawing 3 is stored in the database accumulating part 260 as information about a contribution document, the reply document to this contribution, and the reply document to a reply.

[0039]As shown in drawing 3, the letter-sent-to-get-a-reply-reply relation about each contribution documents A-Z and a reply document is expressed in the series chart by the tree structure. For example, as a reply to the contribution document A, there are the four reply documents A-1-4, and there is two reply document A-1-a-b as a reply document to the reply document A-1. The date, a message number, an implementor name, etc. with which each document was written in the electronic bulletin board are appended to the series chart near each contribution document A-Z and the reply document (not shown). The text of each contribution documents A-Z and a reply document is linked with a series chart, and is accumulated in the database accumulating part 260.

[0040]A two-dot chain line showed an example of a thread mentioned above to drawing 3. That is, reply document A-1-a located in the 3rd layer that is the bottom of the heap of a tree structure to which this contribution document A belongs from the contribution document A of 1 located in the 1st layer that is the top layer of a tree structure is connected due to letter-sent-to-get-a-reply-reply 1 via the reply document A-1 located in the 2nd layer. A settlement (getting it blocked the contribution document A, the reply document A-1, reply document A-1-a) of all the documents connected by such a letter-sent-to-get-a-reply-reply relation serves as the thread TH. As long as it is expressed with a series chart of drawing 3 to drawing 3, "The contribution document A, the reply document A-1, reply document A-1-a", "The contribution document A, the reply document A-1, reply document A-1-b", "the contribution document A, the reply document A-2", The seven threads TH "the contribution document A, the reply document A-3, reply document A-3-a", "the contribution document A and the reply document A-4", "the contribution document B, the reply document B-1 and reply document B-1-a", and "the contribution document Z, the reply document Z-1 and reply document Z-1-a" exist.

[0041]In ROM230 of the database server 200, the program which described the processing (henceforth series chart creation processing) which creates the above-mentioned series chart is stored. This program functions as a sorting means which classifies many documents for every thread as a document group, when CPU220 performs. It is shown in drawing 4 by making the contents and procedure of this series chart creation processing into a series chart creation manipulation routine. This routine is processing which CPU220 performs, when a contribution document and a reply document are written in. Starting of this routine will perform processing which specifies the document first made into the address of the written-in document in a series chart (Step S100). Next, processing which adds some written-in documents (for example, title etc.) is performed, relating with the document specified as the lower layer of the specified document (Step S120). Next, the contribution document and reply document in a series chart are re(Step S140) classified for every thread, the series chart after a classification is updated and memorized to the database accumulating part 260 (Step S160), and this routine is ended.

[0042]In ROM230 of the database server 200. A summary of a contribution document written in an electronic bulletin board other than the above-mentioned series chart creation processing and a reply document is created, A program the contents of processing (henceforth summary creation and display processing) which displays this created summary on a display of the personal computer 900 with a series chart were described to be is stored. It is shown in drawing 5 by making contents and procedure of this summary creation and display processing into summary creation and a display-processing routine. This routine is processing which CPU220 performs, when execution instruction of a purport that a summary of a document written in an electronic bulletin board is created is made.

[0043]Starting of this routine will perform ** thread specific processing (Step S200) first. Thread specific processing is processing which specifies the thread TH of 1 which becomes an object which creates a summary out of a series chart. Henceforth, ** effective text specific processing (Step S210), ** significant sentence extracting processing (Step S220), ** summary creation processing (Step S240), and ** title creation processing (Step S260) are performed for this thread TH of specified 1. Details of these processings are mentioned later. When processing of these steps S220-S260 is completed about all the threads TH, (Step S270) and ** indicative-data output process (Step S280) mentioned later are performed, and this routine is ended.

[0044]Hereafter, the contents of ** effective text specific processing, ** significant sentence extracting processing, ** summary creation processing, ** title creation processing, and the ** indicative-data output process are explained in order. Although the segmentation processing after whole evaluation is included in ** significant sentence extracting processing and sentence standardization processing is included in ** summary creation processing, it combines with ** significant sentence extracting processing and ** summary creation processing, and these processings are also explained.

[0045]** Effective text specific processing (Step S210 of drawing 5)

Effective text specific processing is processing which specifies the text (henceforth the effective text) which extracted expression which cannot constitute the summary of a document from the text of each document which belongs in this thread TH about the thread TH specified by thread specific processing, and excepted extracted expression. In the case of the thread TH enclosed with the two-dot chain line by drawing 3, the text of the contribution document A is specifically read, a thin expression of a relation is excepted from the text to specification of the theme of documents, such as a greeting sentence and a contributor's signature, and the text after deletion is made into the effective text. Also with the reply document A-1 or reply document A-1-a, the text is read and the same processing as the contribution document A is performed. However, in the reply document A-1 or reply document A-1-a, it deletes from the text also about the part which quoted the text of the contribution document A, and the text after deletion is made into the effective text. Such the effective text is memorized temporarily in the predetermined region of RAM240.

[0046]The example of each document (the contribution document A, the reply document A-1, and reply document A-1-a) belonging to the thread TH enclosed with the two-dot chain line by drawing 3 is shown in drawing 6. The contribution document A is a document which was drawn up by the first and written in the electronic bulletin board, and is a document which the reply document A-1 was drawn up by the second as a reply to the contribution document A, and was written in the electronic bulletin board. Reply document A-1-a is the document which was drawn up by the shell as a reply to the reply document A-1, and was written in the electronic

bulletin board. As shown in drawing 6, four sentences, six sentences, and five sentences are contained in the reply document A-1, the reply document A-1, and reply document A-1-a, respectively, but. The finite greeting sentence of "being always indebted" of the 1st sentence of the reply document A-1 is deleted, and the five remaining sentences are made the effective text by execution of the above-mentioned effective text specific processing. in order to make it easy to know next explanation and to obtain, the statement number (1-14) of the sentence made into the effective text is shown in the beginning of a sentence of each sentence in drawing 6.

[0047]** significant sentence extracting processing (Step S220 of drawing 5)

Significant sentence extracting processing is processing which extracts the sentence which has predetermined correlation in each document as a significant sentence from the effective text of each document specified by effective text specific processing. In consideration of the contents expressed with this example in significant sentence extracting processing by the full-effective text which belongs in the thread TH of 1, as a sentence which has predetermined correlation in each document, Processing (henceforth the segmentation processing after whole evaluation) which starts the sentence of the meaning nearest to the contents expressed by the full-effective text from the text of each document is performed. The sentence started by execution of this segmentation processing after whole evaluation is extracted as a significant sentence.

[0048]Although various techniques could be considered, the following technique was adopted as the extracting processing of a significant sentence in this embodiment. The contents and the procedure of significant sentence extracting processing are shown in drawing 7 as a significant sentence extracting processing routine. This routine is started after the end of effective text specific processing. Starting of this routine will perform processing which distinguishes first the kind of each thread TH which an electronic bulletin board has (Step S300). It is for changing the extraction technique of the significant sentence from each document contained in the thread TH according to the kind of thread TH. (the actual contents of the document in a thread). [whether as a kind of this thread, it is developed by the theme which differed from the theme of the first contribution document with each document in a thread, for example, and] Whether it is a thing of the character whether it is a thread of a Q&A system, and in which a letter sent to get a reply and a reply are frequently repeated like a chat in (the form of a thread), and a thread (The character of a thread, a special feature), or the length of one document in a thread is in which tendency of a long sentence or a short sentence (tendency of the document in a thread) -- etc. -- it can think. Suppose that the kind of thread TH is distinguished to the asked type of a question, two or more subject type, or the interactive mode of three in this example.

[0049]Distinction of the kind of thread TH can be performed in the following ways. Hereafter,

the thread TH enclosed with the two-dot chain line by drawing 3 is explained as an example. First, when the effective text of the contribution document A which is the first document that provides the theme is taken out and "the question of -" and expression of "- is not known" are in the effective text, this thread TH is distinguished as it is asked type of a question. In not being applied to the asked type of a question, the full-effective text (effective text of the contribution document A, the reply document A-1, and reply document A-1-a) in the thread TH is taken out, and it judges whether two or more themes are within the limits of [fixed] the full-effective text. When this thread TH is distinguished as it is two or more subject type when there are two or more themes, and two or more themes cannot be found, this thread TH is distinguished as it is an interactive mode. Although the judgment of the existence of two or more themes can be performed by changing the full-effective text into vectorial representation using the TFIDF method, detailed explanation is omitted here.

[0050]In this way, after distinguishing the kind of thread TH, processing which extracts a significant sentence according to the kind of thread TH is performed (Step S320). It attaches in detail [this processing], and divides and explains to the asked type of a question, two or more subject type, and an interactive mode hereafter.

[0051]The extracting processing technique of a significant sentence in case the kind of thread TH is asked type of a question is shown in the flow chart of drawing 8. In being asked type of a question, the first document that provides the theme extracts first a sentence including "the question of -", and expression of "- is not known" from the effective text of ** (the example of drawing 3 the contribution document A), and let this be a significant sentence of the first document (Step S400, S410). Next, the first document and other documents (the example of drawing 3 reply document A-1-a) drawn up by the person same in the same thread TH as the first document are looked for, and a sentence including expression of "it succeeded" or "it having solved", "understanding", etc. is extracted from the effective text of other documents. The extracted sentence turns into a significant sentence of the first document and other documents drawn up by the same person (Step S420, S430). Then, the sentence in which the solution to a question is contained in sentence order including the expression of "having succeeded" of the effective text of a document besides the above is taken out, and the word contained in this sentence is started as a word which shows solution. Next, the first document extracts the sentence in which the word which shows the above-mentioned solution is contained from the effective text of others' document (the example drawing 3's reply document A-1) drawn up by different person. The extracted sentence turns into a significant sentence of others' document (Steps S440-S460).

[0052]Thus, in this example, when the kind of thread TH is asked type of a question. The sentence which has predetermined correlation in this document from each document in the thread TH (in this example.) The standard which starts the sentence of the meaning nearest to

the contents expressed by the full-effective text is changed according to the kind (the example of drawing 3 the contribution document A, the reply document A-1, reply document A-1-a) of document belonging to this document group.

[0053]The extracting processing technique of a significant sentence in case the kinds of thread TH are two or more subject type and an interactive mode is shown in the flow chart of drawing 9. In the case of a subject type and an interactive mode, two or more processings which take out first the effective text (henceforth the full-effective text) of the whole sentence document which belongs in the thread TH of 1 are performed (Step S500).

[0054]In this example, it is considered that the thread TH which has the example shown in drawing 6 was an interactive mode, and the extracting processing of the significant sentence shown in drawing 9 is applied. In the case of the example shown in drawing 6, a total of 14 sentences of the effective text (sentence of the statement numbers 5-9 shown in drawing 6) of the effective text (sentence of the statement numbers 1-4 shown in drawing 6) of the contribution document A and the reply document A-1 and the effective text (sentence of the statement numbers 10-14 shown in drawing 6) of reply document A-1-a are taken out as the full-effective text.

[0055]In a two or more subject type case, in processing of Step S500, the text which belongs to a certain theme of 1 among the full-effective texts, and the text belonging to other themes are taken out independently, make each taken-out text the full-effective text, and processing not more than step S510 is performed.

[0056]Next, a word which is a noun is extracted from the full-effective text, and an extracted word performs processing which searches for a grade which inclines and occurs frequently in each sentence which constitutes the full-effective text (Step S510).

[0057]The grade in which a word inclines and occurs frequently can evaluate the number of times to which the word appears within the full-effective text by a value normalized with the number of sentences which constitute the full-effective text. This is known, for example as TFIDF. TFIDF is defined by the following formula. Each sentence in which db is a mass of target text data (here full-effective text), and d constitutes the full-effective text from the following formulas, and t are taken as a word contained in this text.

[0058]

$$TFIDF(d, t) = TF(d, t) \times IDF(t) \quad -- (1)$$

however, : -- TF (d, t) depends the number of times and ldf to which a word "t" appears in each sentence d on a following formula (2).

$$IDF(t) = \text{LOGe} \{DB(db)/f(t, db)\} \quad -- (2)$$

the number of sentences with which, as for the value 14 and f (t, db), a word "t" appears in the full-effective text here in the number of sentences with which DB (db) constitutes the full-effective text, therefore an example which showed drawing 6 -- it comes out.

[0059]In an example shown in drawing 6, drawing 10 shows a result of having counted the frequency of occurrence of a word which is the noun contained in this example. In this way, after asking for the frequency of occurrence of a word which is a noun, in this example, the frequency of occurrence calculated a TFIDF value about two or more words based on an upper type. In this way, a TFIDF value of each called-for word is shown in drawing 11.

[0060]Then, processing which adds together TFIDF of the word which is a noun which constitutes each sentence is performed (Step S520). The result of this total processing was shown in the column of the total value of drawing 11.

[0061]Next, the size of the total value of each sentence is judged for every effective text of each document contained in the thread TH, and processing which extracts the sentence which is two with the largest total value is performed (Step S530). As with an asterisk shows to drawing 11, in the case of the example of drawing 6, the sentence of the statement numbers 7 and 8 from the reply document A-1 and the sentence of the statement numbers 11 and 14 from reply document A-1-a are extracted for the sentence of the statement numbers 1 and 2 from the contribution document A, respectively.

[0062]Processing to the above steps S510-S530 is equivalent to the segmentation processing after whole evaluation mentioned above. According to this segmentation processing after whole evaluation, if the frequency of occurrence within each effective text is not high, the TFIDF value of a word does not become high. On the other hand, in the case of a word (in for example, the in "things" and the "case") which appears in the full-effective text uniformly, since $f(t, db)$ becomes a big value, a TFIDF value becomes with a value small after all. That is, it becomes a big value as the sentence in which the word which appears by high frequency in each sentence which constitutes each effective text exists, and moreover, it turns into a value with a high total value, so that a TFIDF value has many such a words.

[0063]As mentioned above, two significant sentences are extracted from each document contained in the thread TH. Then, the extracted significant sentence is memorized to the predetermined region of RAM240, this (Step S340 of drawing 7) routine is ended, and it shifts to summary creation processing (Step S240 of drawing 5).

[0064]** Summary creation processing (Step S240 of drawing 5)

They are summary creation processing, next the processing which creates a summary of the thread TH based on a significant sentence extracted by significant sentence extracting processing. In an example of drawing 3, a summary of the thread TH which consists of the contribution document A, the reply document A-1, and reply document A-1-a is created. Creation of this summary is created by specifically combining a significant sentence extracted from each document in a thread based on a predetermined algorithm. For example, the thread TH enclosed with a two-dot chain line is asked type of a question, the contribution document A includes a question in drawing 3, and the reply document A-1 includes a proposal of solution in

it, What is necessary is to enlarge weight which combines a significant sentence extracted from the contribution document A, and a significant sentence extracted from reply document A-1-a, and just to create a summary, in including a result that reply document A-1-a was solved. It is because it is thought that a sentence "the point in question -- was solved" makes a visitor produce interest generally.

[0065]In this example, a summary is created by performing processing (henceforth sentence standardization processing) which rewrites a significant sentence extracted by significant sentence by significant sentence extracting processing to a standard expression. Specifically in sentence standardization processing, short-sentence-izing of a redundant expression, substitution to other independent words of an independent word, correction of direction for use of an attached word, substitution from a dialect to the standard language, replacement to a character defined beforehand, declared unification of a shake, etc. are processed. If it divides roughly into processing of standardization, standardization of a character, declared unification of a shake, unification of an independent word, etc. can be considered. Although explanation is omitted about these processings, it is as follows when rough illustration is mentioned.

[0066]If standardization of a character is illustrated, replacement of (A) parenthesis: "" and "" will be performed, (C) general [replacing quotation-marks: ""and"" etc.] sign : (B) About various signs (for example, ":", "!", etc.). (D) katakana, such as replacing half width/full width, and an alphanumeric character : Replacement of full width/half width, and a capital letter/small letter is performed, (E) period and punctuation-marks: -- a period and punctuation marks -- "-- " -- "." -- ***** , such as transposing *****: "Queen = Elizabeth" of the (F) name to "Queen Elizabeth", such as unifying.

[0067]A shake of a notation means ambiguity of a notation in Japanese, and allowable width. For example, a shake of the example of shake:of ** macron, Windows, Windows, and ** declensional kana ending : An example, the example of shake:of the shake:angel of the example of shake:of a kana notation of sales, sales, the example of shake:of ** contracted sound notation, a wizard, a wizard, and ** compound, sales, sales, and ** loanword transcription, an angel, and ** iteration statement character, and a right right grand one -- **** can be mentioned fairly.

[0068]furthermore -- as unification processing of an independent word -- (b) modifier: -- **** , such as (**) Verb: telling about, connecting, and notifying a (**) noun:personal computer -- it is uncanny, the first and high -- a personal computer, PC, etc., can be illustrated.

[0069]Criterion data and a rule for performing such processing are stored in a dictionary in the hard disk 270 of the database server 200. If replacement to a character defined beforehand is taken for an example, the word "PC" will be transposed to the word a "personal computer."

Thereby, a document in the thread TH of 1 and a term of a significant sentence can be unified.

[0070]** Title creation processing (Step S260 of drawing 5)

Title creation processing is processing which creates a title of a thread based on a significant sentence extracted by significant sentence extracting processing. In an example of drawing 3, a title of the thread TH which consists of the contribution document A, the reply document A-1, and reply document A-1-a is created. A title is created by analyzing grammar information on a significant sentence, specifically starting a required word, and complementing an attached word.

[0071]** indicative-data output process (Step S280 of drawing 5)

An indicative-data output process is processing which outputs the summary created [above-mentioned], a title, and a series chart as an indicative data. This processing is performed when it is judged that the summary and the title were created about all the threads TH of the electronic bulletin board in processing of Step S270 of drawing 5.

[0072]An example of the data image outputted in an indicative-data output process is shown in drawing 12. The data image based on the series chart shown in drawing 3 is expressed with drawing 12. This data image is contents displayed on the preview screen of the display of the personal computer 900 the first stage, when an indicative data reaches the personal computer 900 of a client via the network 100 from the database server 200.

[0073]As shown in drawing 12, the summary and the title are displayed on the preview screen for every thread. The thread name of the thread TH which the thread name is attached based on the reply document located in the bottom of the heap of a series chart, for example, was enclosed with the two-dot chain line by drawing 3 is set to "A-1-a" which is a reply document name of the bottom of the heap. In this example, since they are collectively displayed by a contribution day (day when the contribution document was written in), an updated date (the latest day when the reply document was written in), and the hierarchy number (the number of the reply documents in a thread), these indication can be given the reference works about selection of subject.

[0074]The list display of the title about each thread and the summary is carried out to the order of a time series. In this example, the updated date shows sequentially from a new thread. If it says by drawing 12, a thread name will be the thread by which the thing of "B-1-a" was updated most these days. Thereby, the newest theme and subject can be accessed easily.

[0075]In a preview screen shown in drawing 12, it is good also as displaying a name of a maker of a contribution document belonging to each thread besides being a title, a summary, etc., or a reply document. If it carries out like this, it will become easy to grasp change of utterance contents by each maker.

[0076]The client can choose a thread of subject which is interested with reference to a title or a summary in a preview screen. In this way, selection of a thread will display all the document information which belongs in a selected thread on a list display screen with a letter-sent-to-get-a-reply-reply relation of a document. A thread name shows drawing 13 contents displayed on a

list display screen, when the thread TH of "A-1-a" is chosen. With this list display screen, the contents of interested subject can be checked in detail.

[0077]The bulletin board display system KS of this example described above, The contents expressed by the whole document belonging to the thread TH are taken into consideration about the thread TH which comprised a contribution document which has letter-sent-to-get-a-reply-reply relations, a reply document to contribution, and a reply document to a reply, The sentence of the meaning nearest to the contents of each document to the whole document belonging to the thread TH is extracted as a significant sentence, the summary of the thread TH is created based on the extracted significant sentence, and the created summary is displayed on a screen. Therefore, it becomes possible to grasp the contents of the subject for every thread TH promptly and exactly, and subject can be selected promptly and easily.

[0078]The bulletin board display system KS of this example distinguishes the kind of thread TH, extracts a significant sentence according to the kind of distinguished thread TH, and displays this significant sentence. Therefore, thematic deployment and the contents of subject can be grasped more correctly.

[0079]The bulletin board display system KS of this example applies a sentence standardization means to the extracted significant sentence, rewrites a significant sentence to a standard expression, and creates a summary. Since the sentence of a level uniform irrespective of the skill about a document preparation person's document preparation will be displayed as a summary of a thread if it carries out like this, it becomes much more easy to grasp the contents of the subject for every document group. For example, when the significant sentence which has a redundant expression is short-sentence-ized, the sentence which the contents tend to grasp in the limited viewing area can be displayed.

[0080]The bulletin board display system KS of this example performs extraction of a significant sentence, or creation of a summary based on the effective text which excepted expression which cannot constitute the summary of a document to the document belonging to each thread TH. Therefore, extraction of a significant sentence or creation of a summary can be performed with efficiently and sufficient accuracy.

[0081]The bulletin board display system KS of this example creates the title of the thread TH using the extraction result of a significant sentence, and displays it with the summary of the thread TH. Therefore, as compared with the case where the titles of each document attached by the maker are enumerated, it becomes easy to grasp the kind of subject of each thread TH.

[0082](2) The 2nd embodiment : there is an E-mail which sends each other's document between an addresser and a specific receiver in addition to the write-in document to an electronic bulletin board in the document which has letter-sent-to-get-a-reply-reply relations. A significant sentence is extracted to a settlement of such an E-mail, and the embodiment which creates a summary is described as an e-mail display system MS. Drawing 14 is an explanatory

view showing the outline composition of e-mail display system MS which is the 2nd embodiment of this invention. As shown in drawing 14, although e-mail display system MS is provided with the bulletin board display system KS and almost common composition, it differs from the bulletin board display system KS in that it does not have the database server 200. Double figures are expressed with drawing 14 about this common composition using the same number as drawing 1 on the mark.

[0083]A program which described processing which arranges a settlement (henceforth an e-mail bunch) of mail transmitted and received between a specific addresser and a specific receiver in order of reception thru/or a transmission date, and creates a time sequence diagram table is stored in ROM of the distributing servers 30 and 31. Execution of this program will create a time sequence diagram table as shown in drawing 15. As shown in drawing 15, Mr. Y who received the mails X1 transmitted to Mr. Y from X persons transmitted the mail Y1 to X persons, and X person who received this mail Y1 has transmitted the mail X2 to Mr. Y. In drawing 15, transmission and reception of mail between such X person and Mr. Y continue 5 times. This has chained mail exchanged between X persons and Mr. Y in order of a time series like "the e-mail X1 -> mail Y1 -> mail X2 -> mail Y1 -> mail X3 -> mail Y3 -> mail X4 -> mail Y4 -> mail X5 -> mail Y5." A total of ten copies of mails chained in this way becomes the e-mail bunch MB. The text of each mail in a time sequence diagram table is linked with a time sequence diagram table, and is accumulated in a hard disk in the distributing server 30 and 31.

[0084]The program the same contents of summary creation and display processing as the 1st above-mentioned embodiment were described to be is stored in ROM of the distributing servers 30 and 31. By executing this program with the command of CPU of the distributing servers 30 and 31, ** thread specific processing, effective text specific processing, ** significant sentence extracting processing, ** summary creation processing, ** title creation processing, ** indicative-data output process, and same processing are performed like the thread TH in the 1st embodiment to an e-mail bunch. As a result, the sentence of the meaning nearest to the contents of the whole mail is extracted from each mail which belongs to an e-mail bunch in consideration of the contents expressed by the whole mail belonging to an e-mail bunch as a significant sentence, and the summary of an e-mail bunch is created based on the extracted significant sentence. The summary of such an e-mail bunch is displayed on X persons' preview screen of the both sides the personal computer 90 and Mr. Y's personal computer 91. Therefore, a client becomes possible [grasping promptly and exactly the contents of the subject for every e-mail bunch], and can select e-mail promptly and easily.

[0085]In the above-mentioned embodiment, although the display of the text of e-mail was taken up as an example, the technique of a such display is not limited to the display of a mail text, and can be applied also to the data which comprised two or more sentences. For

example, a significant sentence can be further extracted out of an abstract sentence for the abstract sentence of a paper database, and it can realize in the mode of displaying this with the title of a paper. Or the real original form voice of extracting and displaying a significant sentence from the description of an electronic chart and reports, such as a newspaper distributed electronically, also exists. It is preferred for the extracted significant sentence it not only to to display on a monitor, but to replace with a display on a monitor or to read out with a sound with a display. From the first, printing with a printer etc. does not interfere, either.

[0086]As mentioned above, although the embodiment of the invention was described, As for the ability to carry out with the form which becomes various within limits which do not deviate from the summary of this invention, it is needless to say that this invention is not limited to such an embodiment at all, and displays a significant sentence in a different mode (for example, different colors) from others, for example etc.

[0087]Although it presupposed that the database server 200 performs series chart creation processing (drawing 4), and summary creation and display processing (drawing 5) in the 1st embodiment, It is good also as composition which stores in equipment other than database server 200 the program which described the contents of series chart creation processing, or summary creation and display processing, and equipment other than database server 200 performs. For example, the composition etc. which the distributing server 300,310,320 and the personal computer 900,910,920 of a client perform can be considered.

[0088]Although it presupposed that the distributing servers 30 and 31 perform series chart creation processing, and summary creation and display processing in the 2nd embodiment, It is good also as composition which stores in the distributing server 30 and equipment other than 31 the program which described the contents of series chart creation processing, or summary creation and display processing, and the distributing server 30 and equipment other than 31 perform. For example, the composition etc. which the personal computers 90 and 91 which read the above-mentioned program perform can be considered.

[0089]In the above-mentioned embodiment, in significant sentence extracting processing, the write-in document to an electronic bulletin board, Although the number of average sentences is about five and Kazufumi determined the number of the sentences to extract as two in consideration of being in a comparatively short tendency (processing of Step S530 of drawing 9), the number of the sentences to extract, According to the length etc. of each document (if it says by drawing 3, they will be the contribution document A, the reply document A-1, and reply document A-1-a) contained in the thread TH, it can set arbitrarily.

[0090]In the above-mentioned embodiment, after extracting a significant sentence, sentence standardization processing was performed to this significant sentence, but before extracting a significant sentence, it is good also as composition which performs sentence standardization processing to the effective texts (text of a contribution document or a reply document, etc.).

[0091]Although summary creation processing (Step S240 of drawing 5) and title creation processing (Step S260 of drawing 5) were performed in the above-mentioned embodiment after an end of significant sentence extracting processing (Step S220 of drawing 5), It is good also as composition which outputs a significant sentence of each document extracted by significant sentence extracting processing by an indicative-data output process (Step S280 of drawing 5) so that a display is possible, without performing this summary creation processing and title creation processing. A sentence of a meaning nearest to the contents of each document to the whole document which belongs to the thread TH also by this composition in consideration of contents expressed by the whole document belonging to a thread TH mail bunch is extracted as a significant sentence, and an extracted significant sentence is displayed on a screen. Therefore, it can grasp correctly only by seeing a significant sentence displayed in deployment of the theme in each thread TH or each e-mail bunch.

[0092]When a significant sentence is extracted from each mail which belongs to an e-mail bunch especially and this significant sentence is arranged to a time series, By displaying on a monitor a significant sentence arranged in this way, or printing out, it becomes possible to grasp the past exchange at a glance, and can utilize effectively in a business scene etc. specific contents (for example, delivery date when both sides have agreed) in the past exchange -- investigation -- when like, it is because it is not necessary to say that each mail belonging to an e-mail bunch is displayed on a monitor one by one, or it prints out, and specific contents are looked for from an e-mail bunch.

[0093]Although some embodiments and modifications of this invention were explained above, this invention is not limited to these embodiments at all, and can be carried out in various modes within limits which do not deviate from a summary of this invention. For example, a relation of two or more documents is not restricted to a relation of a letter-sent-to-get-a-reply-reply, and even if they are two or more documents, a comment, a diary, a report that the same maker drew up, for example, it does not interfere. A portion which performs a document display of this invention may be placed on a server which is accumulating two or more documents, and may be set to a client side which peruses two or more documents. Or it is also possible to place in the middle of these client/server architectures as a server for exclusive use. A program for constituting these systems, It is good also as what builds at least some document display systems which it could be dealt with with a form of media, such as CD-ROM, and what places a program on a server and is used downloaded this, and were mentioned above on their computer.

[Translation done.]

* NOTICES *

JP0 and INPIT are not responsible for any damages caused by the use of this translation.

- 1.This document has been translated by computer. So the translation may not reflect the original precisely.
- 2.**** shows the word which can not be translated.
- 3.In the drawings, any words are not translated.

TECHNICAL PROBLEM

[Problem to be solved by the invention]However, in signal transduction on the conventional network. In order it is difficult to grasp correctly a summary of each document which has a predetermined relation at a glance and to grasp a summary of each document correctly, He had to access the text of all the documents in an order from a document located in the top layer, and had to understand the contents of the text, and there was a problem that time and a labor will be applied to grasp of a summary of an exact document. When there were many document numbers especially displayed at once to a client, the above-mentioned problem was more remarkable.

[0008]For example, when it was an electronic bulletin board in which a series chart containing many documents is displayed, it was difficult to grasp an exact summary of a document from a title of each document displayed into a series chart. It is because it is hard to say that it is expressing an exact summary of a document in order that various addressers thru/or reply persons may indicate a title of each document in their different, free style. In particular, in a reply document by a reply person, in such a case, he was not able to understand a summary of a reply document at all from a title so that what added a loess display like "Re:" to the beginning of a title of a contribution document may be automatically made into a title. A summary of a document has not been grasped when it is indicated that it is not related to a summary of a document at the beginning, even if it wrote together a beginning portion of the text of each document on a series chart.

[0009]When many mails are received at once in a cellular phone etc., a majority of some of each mails (for example, beginning portions of a title or the text) are displayed on a display the first stage, but. In order to grasp the summary in which it is difficult and many mails are exact for the same Reason as the above-mentioned electronic bulletin board, grasping the exact summary of e-mail from the title of the mail displayed this first stage a receiver, After the initial display, operation for displaying a mail text further was performed about all the mails, and the

mail text had to be read, scrolling the screen where the mail text was displayed.

[0010]Also in the case of the E-mail etc. which send each other's document between an addresser and a specific receiver, transmitting mail and reception mail have a letter-sent-to-get-a-reply-reply relation, and it has a series of relation as a document with which the theme is common. About the technique grasped correctly at a glance, it was not proposed at all in the former, without accessing the summary of a series of mails which have such a letter-sent-to-get-a-reply-reply relation at each mail text.

[0011]This invention solved this problem, about a series of documents which have letter-sent-to-get-a-reply-reply relations, enabled the grasp of the summary of each document promptly and exactly, and took the following composition for the purpose of enabling it to select document information simply and efficiently.

[Translation done.]

* NOTICES *

JPO and INPIT are not responsible for any damages caused by the use of this translation.

- 1.This document has been translated by computer. So the translation may not reflect the original precisely.
- 2.**** shows the word which can not be translated.
- 3.In the drawings, any words are not translated.

DESCRIPTION OF DRAWINGS

[Brief Description of the Drawings]

[Drawing 1]It is an explanatory view showing the outline composition of the bulletin board display system KS which is the 1st embodiment of this invention.

[Drawing 2]It is an explanatory view showing the structure of the database server 200.

[Drawing 3]It is an explanatory view showing the series chart of the document currently written in the electronic bulletin board.

[Drawing 4]It is a flow chart which shows a series chart creation manipulation routine.

[Drawing 5]It is a flow chart which shows summary creation and a display-processing routine.

[Drawing 6]It is an explanatory view showing the example of each document belonging to the thread TH enclosed with the two-dot chain line by drawing 3.

[Drawing 7]It is a flow chart which shows a significant sentence extracting processing routine.

[Drawing 8]It is a flow chart which shows the extracting processing of a significant sentence in case the kind of thread TH is asked type of a question.

[Drawing 9]It is a flow chart which shows the extracting processing of a significant sentence in case the kinds of thread TH are two or more subject type and an interactive mode.

[Drawing 10]It is an explanatory view showing the example computation which calculated the TFIDF value of each word using the example of drawing 9.

[Drawing 11]It is an explanatory view which illustrates the total value of the sentence unit of the TFIDF value in the example of drawing 9.

[Drawing 12]It is an explanatory view showing the contents displayed on the preview screen of the display of the personal computer 900 after the output of an indicative data the first stage.

[Drawing 13]It is an explanatory view showing an example of a list display screen.

[Drawing 14]It is an explanatory view showing the outline composition of e-mail display system MS as other embodiments.

[Drawing 15]It is an explanatory view showing the time sequence diagram table of the mail

created by the distributing servers 30 and 31.

[Explanations of letters or numerals]

10 -- Network

30, 31 -- Distributing server

90, 91 -- Personal computer

100 -- Network

200 -- Database server

210 -- Network interface

220 -- CPU

230 -- ROM

240 -- RAM

250 -- Timer

260 -- Database accumulating part

270 -- Hard disk

300,310,320 -- Distributing server

800 -- Personal computer

900,910,920 -- Personal computer

KS -- Bulletin board display system

MS -- E-mail display system

MB -- E-mail bunch

TH -- Thread

[Translation done.]

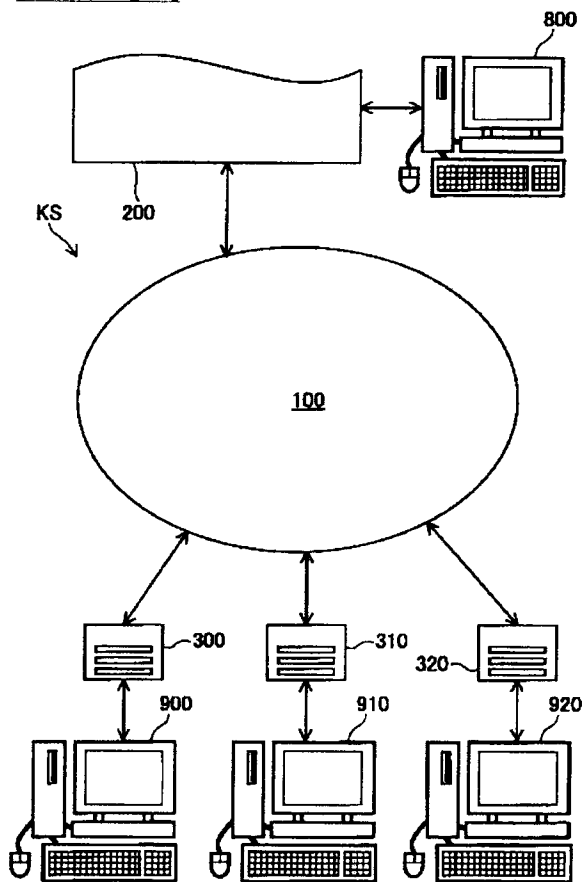
* NOTICES *

JP0 and INPIT are not responsible for any damages caused by the use of this translation.

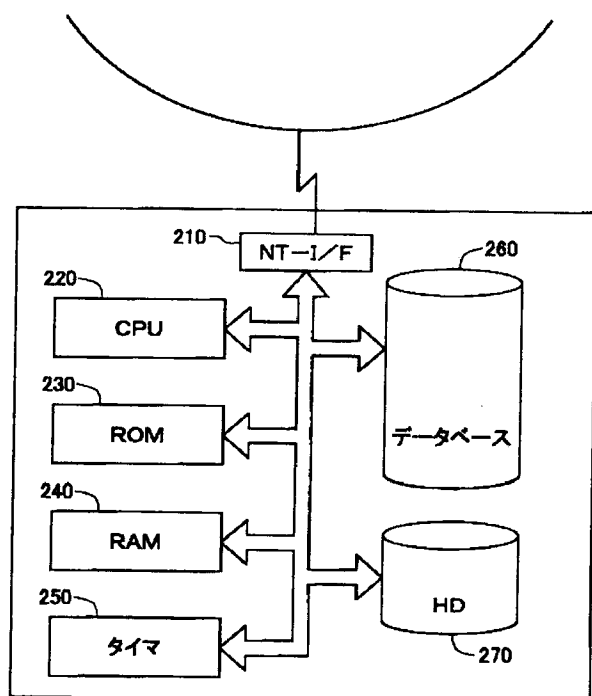
- 1.This document has been translated by computer. So the translation may not reflect the original precisely.
- 2.**** shows the word which can not be translated.
- 3.In the drawings, any words are not translated.

DRAWINGS

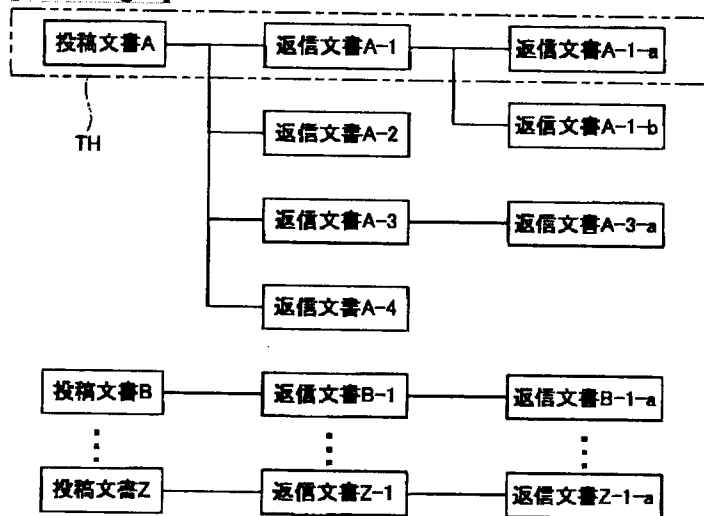
[Drawing 1]



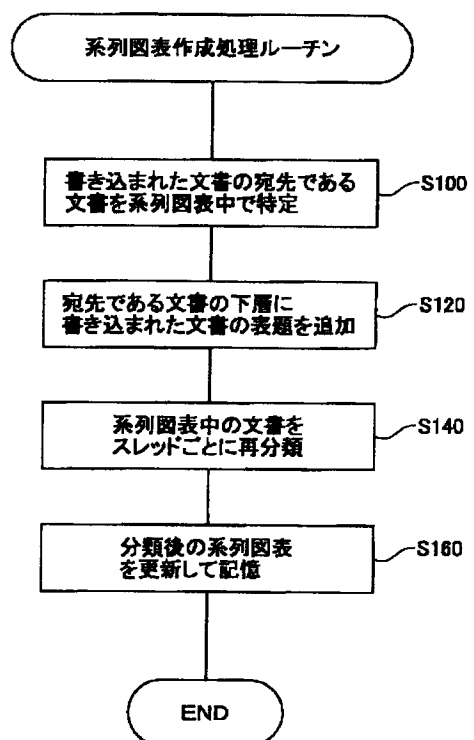
[Drawing 2]



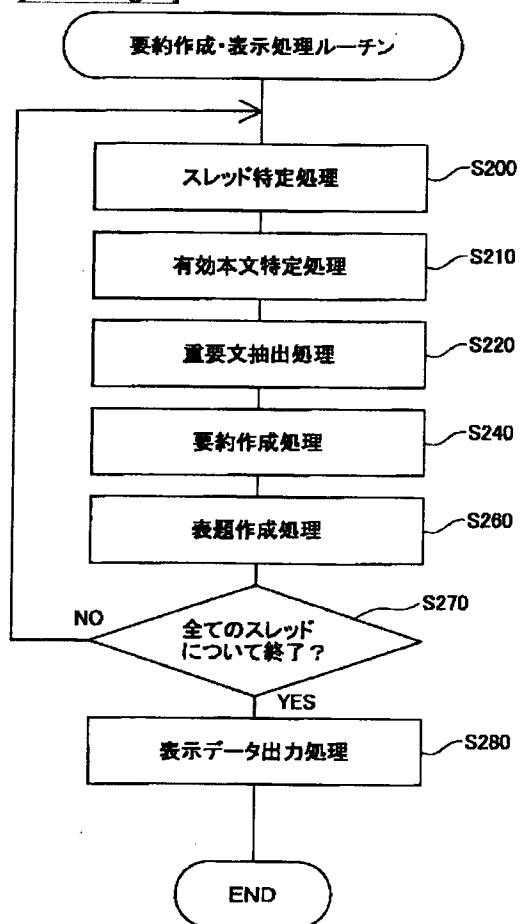
[Drawing 3]



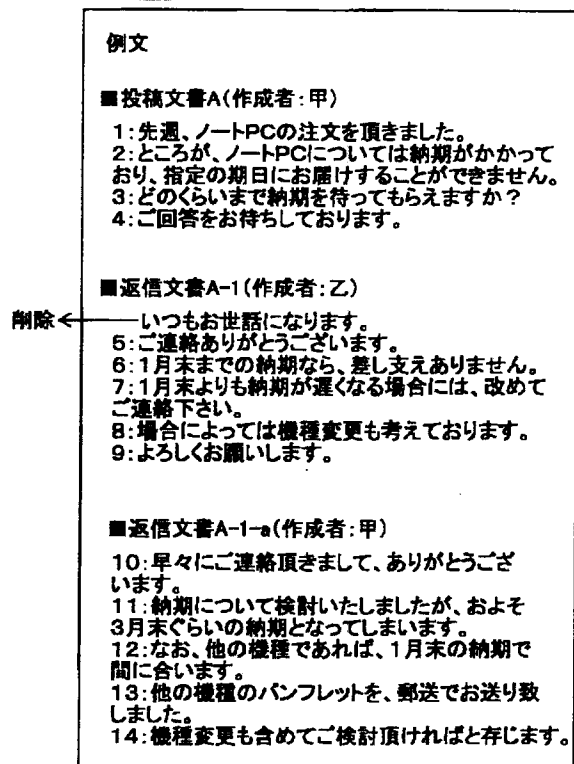
[Drawing 4]



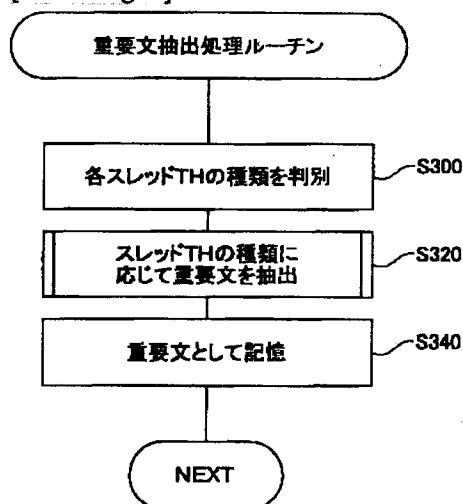
[Drawing 5]



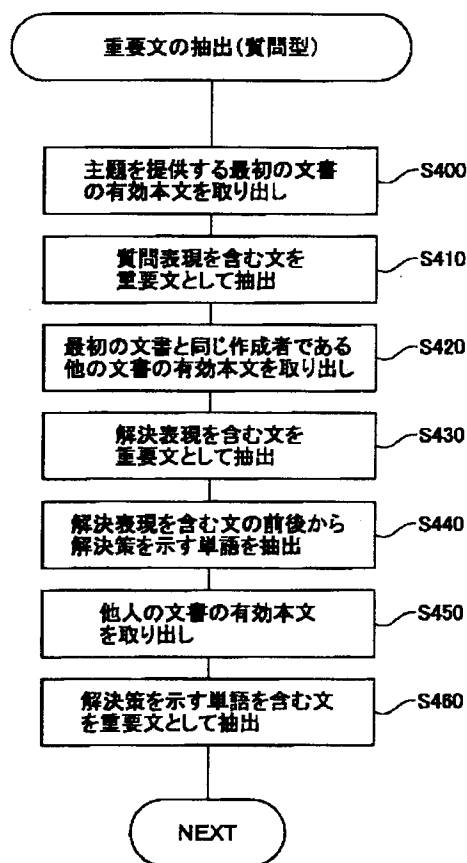
[Drawing 6]



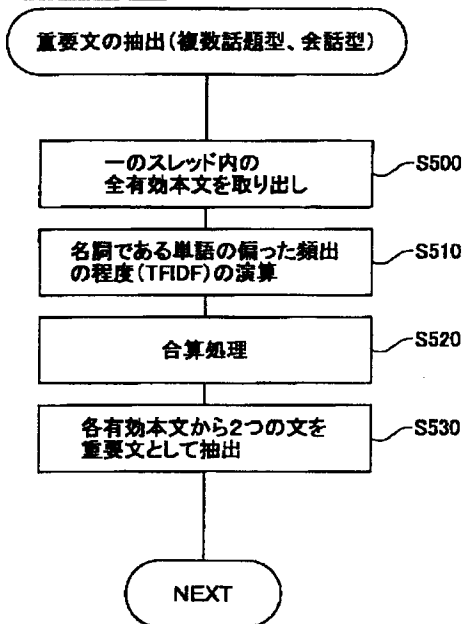
[Drawing 7]



[Drawing 8]



[Drawing 9]



[Drawing 10]

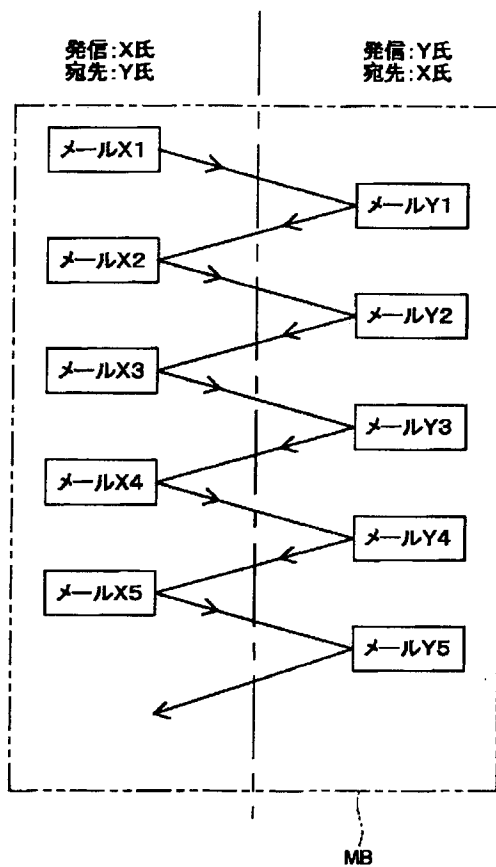
	文番号	先通	ノート	パソコン	注文	納期	指定	期日	回答	連絡	機種	変更	パンフレット
投稿 文書A	1	1	1	1	1								
	2		1	1		1	1	1					
	3					1							
	4								1				
	5									1			
返信文書 A-1	6					1							
	7					1				1			
	8										1	1	
	9												
返信文書 A-1-a	10									1			
	11					2							
	12										1		
	13										1		1
	14										1	1	
	合計	1	2	2	1	6	1	1	1	3	4	2	1

[Drawing 11]

出現頻度2以上の単語についての演算結果

	文番号	合算値	ノート	パソコン	納期	連絡	機種	変更
投稿 文書A	★ 1	3.892	1.946	1.946				
	★ 2	4.739	1.946	1.946	0.847			
	3	0.847			0.847			
	4	1.946						
返信文書 A-1	5	1.540				1.540		
	6	0.847			0.847			
	★ 7	2.387			0.847	1.540		
	★ 8	3.199					1.253	1.946
	9	0						
返信文書 A-1-a	10	1.540				1.540		
	★ 11	1.695			1.695			
	12	1.253					1.253	
	13	1.253					1.253	
	★ 14	3.199					1.253	1.946

[Drawing 15]



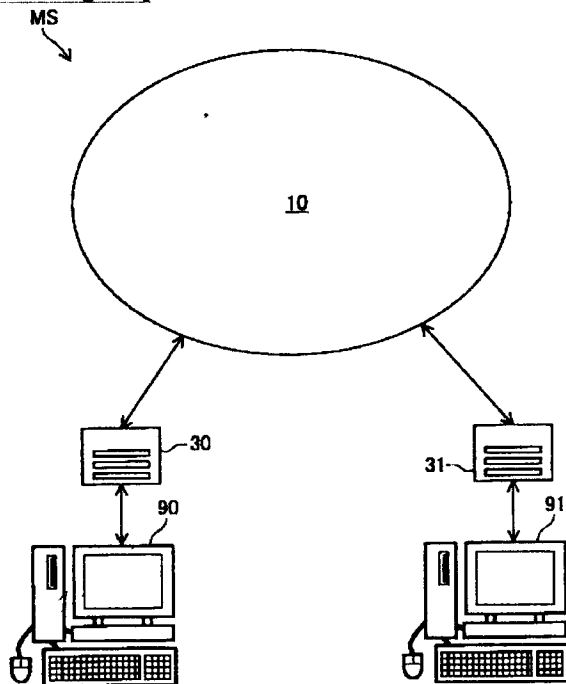
[Drawing 12]

スレッド名	表題	要約	投稿日	更新日	階層数
B-1-a	11/28	12/5	3
A-1-a	ノートパソコンの 納期について	ノートパソコンは納期がかかっており、 機種の変更が必要な場合がある。	11/25	11/30	3
A-4	11/25	11/28	2
A-1-b	11/25	11/27	3

[Drawing 13]

スレッド名:A-1-a
<p>作成者: 甲 書込日: 2001/11/25</p> <p>先週、ノートPCの注文を頂きました。ところが、ノートPCについては納期がかかっており、指定の期日にお届けすることができません。どのくらいまで納期を待ってもらえますか？ご回答をお待ちしております。</p>
<p>作成者: 乙 書込日: 2001/11/26</p> <p>いつもお世話になります。ご連絡ありがとうございます。 1月末までの納期なら、差し支えありません。 1月末よりも納期が遅くなる場合には、改めてご連絡下さい。 場合によっては機種変更も考えております。よろしくお願いします。</p>
<p>作成者: 甲 書込日: 2001/11/30</p> <p>早々にご連絡頂きまして、ありがとうございます。納期について検討いたしましたが、およそ3月末ぐらいの納期となってしまいます。 なお、他の機種であれば、1月末の納期で間に合います。 他の機種のパンフレットを、郵送でお送り致しました。 機種変更も含めてご検討頂ければと存じます。</p>

[Drawing 14]



[Translation done.]

(19)日本国特許庁 (J P)

(12) 公開特許公報 (A)

(11)特許出願公開番号

特開2003-108278

(P2003-108278A)

(43)公開日 平成15年4月11日(2003.4.11)

(51)Int.Cl.

G 0 6 F 3/00

識別記号

6 5 1

F I

G 0 6 F 3/00

ナコード*(参考)

6 5 1 B 5 E 5 0 1

審査請求 未請求 請求項の数17 O L (全 17 頁)

(21)出願番号 特願2001-303205(P2001-303205)

(22)出願日 平成13年9月28日(2001.9.28)

(71)出願人 000002369

セイコーエプソン株式会社

東京都新宿区西新宿2丁目4番1号

(72)発明者 河西 廣雄

長野県諏訪市大和三丁目3番5号 セイコ

ーエプソン株式会社内

(72)発明者 山田 孝司

長野県諏訪市大和三丁目3番5号 セイコ

ーエプソン株式会社内

(74)代理人 110000028

特許業務法人 明成国際特許事務所

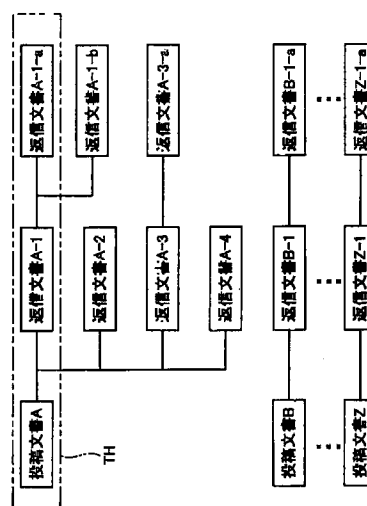
最終頁に続く

(54)【発明の名称】 文書表示システム、受付装置、文書表示方法、および記録媒体

(57)【要約】

【課題】 往信-返信関係を有する一連の文書について、各文書の要旨を迅速かつ的確に把握可能とし、文書情報の取捨選択を簡単かつ効率的に行なえるようにすることを目的とする。

【解決手段】 掲示板表示システムKSでは、電子掲示板の各スレッドTHにおいて、スレッドTHに属する各文書(投稿文書A、返信文書A-1および返信文書A-1-a)から重要文を抽出し、各重要文からスレッドTHの要約を作成して画面に表示する。上記の重要文の抽出においては、一のスレッドTH内に属する全文書の本文によって表わされる内容を考慮し、全文書の本文によって表わされる内容に最も近い意味の文を各文書の本文から重要文として抽出する。



【特許請求の範囲】

【請求項1】 所定の関係を有する複数の文書の少なくとも一部を画面上に閲覧可能に表示する文書表示システムであって、前記複数の文書からひとまとまりの文書群を抽出する文書群抽出手段と、前記文書群に属する文書全体によって表わされる内容を考慮して、前記文書群に属する各文書から該文書に所定の相関を有する文を重要文として抽出する重要文抽出手段と、該重要文抽出手段により抽出された重要文を、前記文書群に属する文書の一部として前記画面上に表示する重要文表示手段とを備えた文書表示システム。

【請求項2】 請求項1に記載の文書表示システムであって前記所定の関係を有する複数の文書は、往信と返信の関係にある文書である文書表示システム。

【請求項3】 請求項1または2に記載の文書表示システムであって、前記文書群の種類を判別する種類判別手段を備え、前記重要文抽出手段は、該種類判別手段により判別された種類に応じて前記重要文を抽出する手段である文書表示システム。

【請求項4】 前記文書群の種類は、質問表現を含む質問型、複数の話題を含む複数話題型、会話文を含む会話型のうち少なくとも一つである請求項3に記載の文書表示システム。

【請求項5】 前記重要文抽出手段は、前記文書群に属する各文書から該文書に所定の相関を有する文を重要文として抽出する基準を該文書群に属する文書の種類に応じて異ならせる相関設定手段を備えた請求項3または4に記載の文書表示システム。

【請求項6】 請求項1ないし5のいずれかに記載の文書表示システムであって、前記重要文抽出手段により抽出された重要文に基づいて前記文書群の要約を作成する要約作成手段を備え、前記重要文表示手段に替えて、該要約作成手段により作成された要約を前記画面上に表示する要約表示手段を備えた文書表示システム。

【請求項7】 前記重要文抽出手段により抽出された重要文を、前記文書群に属する文書の一部として前記画面上に表示する重要文表示手段を備えた請求項6に記載の文書表示システム。

【請求項8】 請求項1ないし請求項7のいずれかに記載の文書表示システムであって、所定の基準に基づいて前記各文書の文を標準的な表現に書き換える文標準化手段を備えた文書表示システム。

【請求項9】 請求項6または7に記載の文書表示システムであって、前記要約作成手段は、少なくとも前記重要文抽出手段により抽出された重要文の表現を標準的な表現に書き換え

る手段を備える文書表示システム。

【請求項10】 請求項1ないし9のいずれかに記載の文書表示システムであって、

所定の基準に基づいて前記文書の中から文書の要旨を構成し得ない表現を除いた文を意味文として抽出する意味文抽出手段を備え、

前記重要文抽出手段または前記要約作成手段は、該意味文を用いて、前記重要文の抽出または前記要約の作成を行なう手段である文書表示システム。

【請求項11】 前記重要文表示手段または要約表示手段は、前記文書群についての重要文または要約を時系列順に一覧表示する手段を備えた請求項1ないし10のいずれかに記載の文書表示システム。

【請求項12】 前記重要文表示手段または要約表示手段は、前記文書群に属する文書の作成者を、前記重要文または前記要約と関連付けて表示する手段を備えた請求項11に記載の文書表示システム。

【請求項13】 請求項1ないし12のいずれかに記載の文書表示システムであって、

前記重要文抽出手段による抽出に基づいて前記文書群の表題を作成する表題作成手段を備え、

前記重要文表示手段または要約表示手段は、該表題作成手段により作成された各文書の表題を、前記重要文または前記要約と関連付けて表示する手段を備えた文書表示システム。

【請求項14】 通信回線を介して文書を受信する受信装置であって、

前記受信した複数の文書からひとまとまりの文書群を抽出する文書群抽出手段と、

該文書群抽出手段により抽出された一の文書群に関し、該文書群に属する各文書から少なくとも本文を取り出す本文取出手段と、

該本文取出手段により取り出された本文全体によって表わされる内容を考慮し、前記文書群に属する各文書から該文書に所定の相関を有する文を重要文として抽出する抽出手段と、

該抽出手段により抽出された前記一の文書群についての重要文を通信回線を介して出力する出力手段とを備えた受信装置。

【請求項15】 請求項14に記載の受信装置であって、

前記抽出手段により抽出された重要文に基づいて前記文書群の要約を作成する作成手段を備え、

前記出力手段に替えて、該作成手段により作成された要約を通信回線を介して出力する要約出力手段を備えた受信装置。

【請求項16】 所定の関係を有する複数の文書の少なくとも一部を画面上に閲覧可能に表示する文書表示方法であって、

前記複数の文書からひとまとまりの文書群を抽出し、

前記文書群に属する文書全体によって表わされる内容を考慮して、前記文書群に属する各文書から該文書に所定の相関を有する文を重要文として抽出し、該抽出された重要文を、前記文書群に属する文書の一部として前記画面上に表示する文書表示方法。

【請求項17】 所定の関係を有する複数の文書の少なくとも一部を画面上に閲覧可能に表示するためのコンピュータプログラムを記録した記録媒体であって、前記複数の文書からひとまとまりの文書群を抽出する工程と、前記文書群に属する文書全体によって表わされる内容を考慮して、前記文書群に属する各文書から該文書に所定の相関を有する文を重要文として抽出する工程と、該抽出された重要文を、前記文書群に属する文書の一部として前記画面上に表示する工程とをコンピュータに実行させるためのプログラムをコンピュータに読み取り可能に記録した記録媒体。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】本発明は、文書表示システムに関し、詳しくは、所定の関係を有する複数の文書から構成された文書群に関し、該文書群に属する文書の少なくとも一部を画面上に閲覧可能に表示する技術に関する。

【0002】

【従来の技術】通信技術の発達した今日では、インターネットやパソコン通信等のネットワーク上で複数人が文字や画像等を用いて情報を伝達し合うことが日常的に行なわれている。このような情報伝達は、ネットワーク上以外での情報伝達と同様に、他人に向けて自発的に情報を発信する者（発信者）および該情報発信者によって発信された情報を受け取る者（受領者）、受け取った情報に対する回答や意見等を情報発信者に返信する者（返信者）が存在することにより成立している。

【0003】一方、ネットワーク上での情報伝達は、発信者と受領者ないし返信者とのやりとりが、データの送受信や受信データの閲覧等のサービスを提供するサーバを介して行なわれるという特徴がある。即ち、発信者ないし返信者は、発信ないし返信しようとする情報をネットワークを介してサーバに送る。情報を受け取ったサーバは、受け取った情報をネットワークを介して受領者に送信したり、受け取った情報を受領者がネットワークを介して閲覧可能に表示したり、前記受領者への送信や受領者による閲覧のために保存したりするのである。なお、サーバに接続して上記のサービスを受ける者（上記の発信者や受領者、返信者）のことを、以下、クライアントという。

【0004】近年では、ネットワーク上での情報伝達の手法として、発信者と特定の受領者との間で文書を送り合うことにより情報伝達を行なう手法（例えば、プッシ

ュ型の電子メール等）以外に、サーバに設けられた電子掲示板（ＢＢＳ）への文書の書き込みにより多数の者の間で情報伝達を行なう手法が用いられている。具体的には、ある主題の文書を他人に発信しようとする発信者は、発信しようとする文書（以下、投稿文書という）の内容を電子掲示板に書き込む。電子掲示板に書き込まれた投稿文書の内容は、ネットワークを介して多数の者（不特定人若しくは特定のグループ全員）に閲覧可能となる。閲覧により投稿文書の内容を受け取った受領者は、この投稿文書に対する返答や意見等を内容とする文書（以下、投稿に対する返信文書という）を電子掲示板に書き込むことができる。書き込まれた投稿に対する返信文書の内容は、投稿文書と同様に多数の者に閲覧可能となる。閲覧により返信文書の内容を受け取った受領者は、この返信文書に対する返答や意見等を内容とする文書（以下、返信に対する返信文書という）を電子掲示板に書き込むことができる。書き込まれた返信に対する返信文書は、投稿文書等と同様に多数の者に閲覧可能となる。

【0005】投稿文書と投稿に対する返信文書は、主題が共通する文書として一連のつながりを有しており、所定の関係を有している。このため、従来の電子掲示板では、往信－返信関係を有する一連の文書を一の文書群とみなし、各文書群に属する文書をツリー構造（枝分かれをした階層構造）に整理していた。ここで、ツリー構造の最上層に位置する一の投稿文書から該投稿文書が属するツリー構造の最下層に位置する返信文書までが一の往信－返信関係で接続されるとき、該一の往信－返信関係で接続される全ての文書のまとまりを、以下、スレッドという。

【0006】従来の電子掲示板では、上記のツリー構造に基づいて各文書群に属する文書の系列図表を作成して、この系列図表を閲覧可能に表示し、文書間の往信－返信関係をクライアントが容易に把握できるようにしていた。また、表示される系列図表中に各文書の表題や各文書の本文の冒頭部分を併記することにより、クライアントが、各文書の本文にアクセスする前に文書の概要を把握できるようにしていた。

【0007】

【発明が解決しようとする課題】しかしながら、従来のネットワーク上での情報伝達では、所定の関係を有する各文書の要旨を一目で正確に把握することは困難であり、各文書の要旨を正確に把握するためには、最上層に位置する文書から順番に全ての文書の本文にアクセスして本文の内容を理解しなければならず、正確な文書の要旨の把握に時間と労力がかかってしまうという問題があった。特に、クライアントに対して一度に表示される文書数が多い場合には、上記の問題はより顕著であった。

【0008】例えば、多数の文書を含む系列図表が表示される電子掲示板の場合には、系列図表中に表示される

各文書の表題から文書の正確な要旨を把握することは難しかった。各文書の表題は、種々の発信者ないし返信者が思い思いの自由なスタイルで記載したものであるため、文書の正確な要旨を表現しているとは言い難いからである。特に、返信者による返信文書においては、投稿文書の表題の冒頭に「Re:」のようなレス表示を付加したものが自動的に表題とされるような場合があり、このような場合には返信文書の要旨を表題から全く理解することができなかった。また、系列図表に各文書の本文の冒頭部分を併記しても、冒頭に文書の要旨とは関係ないことが記載されている場合には、文書の要旨を把握することはできなかった。

【0009】また、携帯電話等において多数のメールを一度に受信した場合には多数の各メールの一部（例えば、表題や本文の冒頭部分）がディスプレイに初期表示されるが、この初期表示されるメールの表題からメールの正確な要旨を把握することは上記電子掲示板と同様の理由により難しく、多数のメールの正確な要旨を把握するために、受領者は、初期表示後に、全てのメールについて、メール本文を更に表示するための操作を行なうと共に、メール本文が表示された画面をスクロール等しながらメール本文を読まなければならなかった。

【0010】また、発信者と特定の受領者との間で文書を送り合う電子メール等の場合にも、送信メールと受信メールとが往信-返信関係にあり、主題が共通する文書として一連のつながりを有している。このような往信-返信関係を有する一連のメールの要旨を、各メール本文にアクセスすることなく、一目で正確に把握する手法については、従来において何ら提案されていなかった。

【0011】本発明は、かかる問題を解決し、往信-返信関係を有する一連の文書について、各文書の要旨を迅速かつ的確に把握可能とし、文書情報の取捨選択を簡単かつ効率的に行なえるようにすることを目的として、以下の構成を採った。

【0012】
【課題を解決するための手段およびその作用・効果】本発明の文書表示システムは、所定の関係を有する複数の文書の少なくとも一部を画面上に閲覧可能に表示する文書表示システムであって、前記複数の文書からひとまとまりの文書群を抽出する文書群抽出手段と、前記文書群に属する文書全体によって表わされる内容を考慮して、前記文書群に属する各文書から該文書に所定の相関を有する文を重要文として抽出する重要文抽出手段と、該重要文抽出手段により抽出された重要文を、前記文書群に属する文書の一部として前記画面上に表示する重要文表示手段とを備えたことを要旨とする。

【0013】上記発明の文書表示システムによれば、重要文抽出手段が、ひとまとまりの文書群に属する文書全体によって表わされる内容を考慮して、文書群に属する各文書から該文書に所定の相関を有する文を重要文とし

て抽出し、抽出された重要文を重要文表示手段が文書の一部として画面上に表示する。従って、文書群における主題の展開を、表示された重要文を見ただけで正確に把握することができる。上記所定の関係を有する複数の文書を、往信と返信の関係にある文書上とすることも好ましい。

【0014】文書群の種類を判別する種類判別手段を備え、該種類判別手段により判別された種類に応じて重要文を抽出する構成としてもよい。こうすれば、文書群の種類に応じた重要文が表示されるので、主題の展開をより正確に把握することができる。

【0015】この場合において、文書群の種類を、質問表現を含む質問型、複数の話題を含む複数話題型、会話文を含む会話型のうち少なくとも一つとすることも好適である。また、重要文抽出手段が、文書群に属する各文書から該文書に所定の相関を有する文を重要文として抽出する基準を該文書群に属する文書の種類に応じて異ならせる相関設定手段を備えることも望ましい。

【0016】重要文抽出手段により抽出された重要文に基づいて文書群の要約を作成する要約作成手段を備え、重要文表示手段に替えて、該要約作成手段により作成された要約を画面上に表示する要約表示手段を備えることも望ましい。こうすれば、文書群ごとの話題の内容を迅速かつ的確に把握することができる。

【0017】こうした要約表示手段と共に、重要文抽出手段により抽出された重要文を文書群に属する文書の一部として前記画面上に表示する重要文表示手段を備える構成としても差し支えない。

【0018】所定の基準に基づいて前記各文書の文を標準的な表現に書き換える文標準化手段を備えることも好適である。また、要約作成手段が、少なくとも前記重要文抽出手段により抽出された重要文の表現を標準的な表現に書き換える手段を備える構成とすれば、文書作成者の文書作成に関する技量に拘らず均一なレベルの文が文書や文書群の要旨として表示されるので、文書群ごとの話題の内容をより一層把握しやすくなる。

【0019】上記の文標準化手段としては、例えば、冗長な表現の短文化や自立語の他の自立語への置換、付属語の用法の訂正、方言から標準語への置換、予め定めた文字への置き換え、表記のゆれの統一等を考えることができる。

【0020】所定の基準に基づいて文書の中から文書の要旨を構成し得ない表現を除いた文を有意味文として抽出する有意味文抽出手段を備え、重要文抽出手段または要約作成手段を、該有意味文を用いて重要文の抽出または要約の作成を行なう手段とすることも好適である。上記の表現抽出手段により抽出される表現としては、例えば、以前の発言の引用箇所や挨拶文、発信者の署名等を考えることができる。こうすれば、重要文の抽出または要約の作成を効率的かつ精度よく行なうことができる。

【0021】重要文表示手段または要約表示手段を、文書群についての重要文または要約を時系列順に一覧表示する手段とすることも好適である。こうすれば、最新の主題や話題に容易にアクセスすることができる。

【0022】重要文表示手段または要約表示手段を、文書群に属する文書の作成者を重要文または要約に関連付けて表示する手段としてもよい。こうすれば、各作成者による発言内容の変化を把握しやすくなる。

【0023】重要文抽出手段による抽出に基づいて文書群の表題を作成する表題作成手段を備え、該表題作成手段により作成された各文書群の表題を重要文または要約に関連付けて表示しても差し支えない。こうすれば、作成者によって付けられた各文書の表題を羅列する場合と比較して、各文書群の話題の種類を把握しやすくなる。

【0024】本発明の受信装置は、通信回線を介して文書を受信する受信装置であって、前記受信した複数の文書からひとまとまりの文書群を抽出する文書群抽出手段と、該文書群抽出手段により抽出された一の文書群に関し、該文書群に属する各文書から少なくとも本文を取り出す本文取出手段と、該本文取出手段により取り出された本文全体によって表わされる内容を考慮し、前記文書群に属する各文書から該文書に所定の相関を有する文を重要文として抽出する抽出手段と、該抽出手段により抽出された前記一の文書群についての重要文を通信回線を介して出力する出力手段とを備えたことを要旨とする。

【0025】上記発明の受信装置によれば、文書群抽出手段が、通信回線を介して受信した複数の文書からひとまとまりの文書群を抽出し、抽出された一の文書群に属する各文書から本文取出手段が少なくとも本文を取り出す。抽出手段は、取り出された本文全体によって表わされる内容を考慮して、文書群に属する各文書から該文書に所定の相関を有する文を重要文として抽出し、抽出された一の文書群についての重要文を出力手段が通信回線を介して出力する。従って、出力された重要文を通信回線を介して得ることにより、一の文書群における主題の展開を正確に把握可能なデータを得ることが可能となる。例えば、このデータを表示や印刷等すれば、文書群における主題の展開を、表示しない印刷された重要文を見ただけで正確に把握することができる。

【0026】なお、上記の文書群形成手段、本文取出手段、抽出手段および出力手段を二以上の装置に分けて設け、これら二以上の装置を組み合わせて上記発明の受信装置を実現することも可能である。

【0027】抽出手段により抽出された重要文に基づいて文書の要約を作成する作成手段を備え、該作成手段により作成された要約を通信回線を介して出力する要約出力手段を出力手段に替えて備えることも望ましい。こうすれば、出力された要約を通信回線を介して得ることにより、文書群ごとの話題の内容を迅速かつ的確に把握可能なデータを得ることが可能となる、データの利用価値

を高めることができる。

【0028】本発明の文書表示方法は、所定の関係を有する複数の文書の少なくとも一部を画面上に閲覧可能に表示する文書表示方法であって、前記複数の文書からひとまとまりの文書群を抽出し、前記文書群に属する文書全体によって表わされる内容を考慮して、前記文書群に属する各文書から該文書に所定の相関を有する文を重要文として抽出し、該抽出された重要文を、前記文書群に属する文書の一部として前記画面上に表示することを要旨とする。

【0029】上記発明の文書表示方法によれば、ひとまとまりの文書群に関し、該文書群に属する文書全体によって表わされる内容を考慮して、各文書から該文書に所定の相関を有する文を重要文として抽出し、抽出された重要文を文書の一部として画面上に表示する。従って、文書群における主題の展開を、表示された重要文を見ただけで正確に把握することができる。

【0030】本発明のコンピュータプログラムを記録した記録媒体は、所定の関係を有する複数の文書の少なくとも一部を画面上に閲覧可能に表示するためのコンピュータプログラムを記録した記録媒体であって、前記複数の文書からひとまとまりの文書群を抽出する工程と、前記文書群に属する文書全体によって表わされる内容を考慮して、前記文書群に属する各文書から該文書に所定の相関を有する文を重要文として抽出する工程と、該抽出された重要文を、前記文書群に属する文書の一部として前記画面上に表示する工程とをコンピュータに実行させるためのプログラムをコンピュータに読み取り可能に記録したことを要旨とする。

【0031】上記発明の記録媒体によれば、記録されたプログラムがコンピュータに読み取られることにより、ひとまとまりの文書群に属する文書全体によって表わされる内容を考慮して、各文書から該文書に所定の相関を有する文を重要文として抽出し、抽出された重要文を文書の一部として画面上に表示する。従って、文書群における主題の展開を、表示された重要文を見ただけで正確に把握することができる。

【0032】なお、上記の記録媒体が備える各工程を二以上の記録媒体に分けて記録し、これら二以上の記録媒体を組み合わせることで上記発明の記録媒体を実現することも可能である。

【0033】なお、記録媒体としては、フレキシブルディスク、CD-ROM、DVD-ROM、半導体メモリ（ROM、PROM、EEPROM、フラッシュメモリ等）など、種々の記録媒体を用いることができる。もとより、インターネットなどのネットワーク上に置かれたサーバにこれらのプログラムを記憶しておき、クライアントのコンピュータにダウンロードして利用することも可能である。

【0034】また、重要文や要約等の表示は、ディスプ

レイや用紙上に表示するものとどまらず、重要文を音声合成などを利用して読み上げることで表示することも可能である。耳から聞いて内容を認識する場合、何が重要かをすぐに判別することは、文字を読み取って判断することより困難なので、重要文を抽出して読み上げる構成は、極めて有用である。携帯電話でメール文書を確認するような場合には、重要文を音声で読み上げれば、携帯電話の小さなディスプレイに制約されることがなく、この点でも有用である。

【0035】

【発明の実施の形態】以下、本発明の実施の形態を実施例に基づいて説明する。

(1) 実施例の構成：はじめに、実施例の構成について図1を用いて説明する。図1は、本発明の第1実施例である掲示板表示システムKSの概略構成を示す説明図である。図1に示すように、掲示板表示システムKSでは、インターネットのような大規模なネットワーク100に、配信サーバ300、310、320・・・およびデータベースサーバ200が接続されている。配信サーバ300には、クライアントのパーソナルコンピュータ900、910、920・・・が電話回線を介して接続されている。また、電子掲示板のウェブページが格納されたサーバ200には、電子掲示板の管理人のパーソナルコンピュータ800が電話回線を介して接続されている。

【0036】パーソナルコンピュータ900から配信サーバ300に対して電子掲示板のウェブページへのアクセス信号が送出されると、配信サーバ300とデータベースサーバ200内の電子掲示板のウェブページとがネットワーク100を介して接続される。これにより、パーソナルコンピュータ900のディスプレイに電子掲示板のウェブページが表示される。この後、クライアントが電子掲示板への文書の掲載を希望する場合には、掲載しようとする文書をパーソナルコンピュータ900を用いて作成し、作成した文書を電子掲示板のアドレスに送る。作成された文書は、配信サーバ300からネットワーク100を介してデータベースサーバ200内の電子掲示板のウェブページに到達する。これにより、クライアントの作成した文書が電子掲示板に書き込まれ、パーソナルコンピュータ900のディスプレイには、書き込み後の電子掲示板が表示される。なお、他のクライアントのパーソナルコンピュータ910、920・・・を用いた場合も、上記パーソナルコンピュータ900と同様に、電子掲示板へのアクセスないし書き込みが可能である。

【0037】データベースサーバ200の構造を図2に示す。データベースサーバ200は、ネットワーク100とのデータのやり取りを制御するネットワークインタフェース(NT-I/F)210、処理を行なうCPU220、処理プログラムや固定的なデータを記憶するR

OM230、ワークエリアとしてのRAM240、時間を管理するタイマ250、後述する各種のデータを蓄積するデータベース蓄積部260、辞書などを記憶しているハードディスク270等を備える。なお、データベース蓄積部260は、実際には、ハードディスクなどの記憶装置に格納されているが、ここでは、説明の都合上、独立の装置として扱うものとする。

【0038】データベース蓄積部260には、電子掲示板に書き込まれている多数の文書が蓄積されている。これらの文書は、前述した往信-返信関係に基づいて整理された状態で蓄積されている。具体的には、データベース蓄積部260には、投稿文書やこの投稿に対する返信文書、返信に対する返信文書に関する情報として、図3に示すような文書の系列図表が格納されている。

【0039】図3に示すように、系列図表には、各投稿文書A～Zおよび返信文書に関する往信-返信関係がツリー構造で表されている。例えば、投稿文書Aに対する返信としては四つの返信文書A-1～4があり、返信文書A-1に対する返信文書としては二つの返信文書A-1-a～bがある。なお、系列図表には、各投稿文書A～Zおよび返信文書の近傍に、各文書が電子掲示板に書き込まれた日付け、メッセージ番号、作成者名等が付記されている(図示せず)。また、各投稿文書A～Zおよび返信文書の本文は、系列図表とリンクされてデータベース蓄積部260に蓄積されている。

【0040】前述したスレッドの一例を、図3に二点鎖線で示した。即ち、ツリー構造の最上層である第1層に位置する一の投稿文書Aから該投稿文書Aが属するツリー構造の最下層である第3層に位置する返信文書A-1-aまでは、第2層に位置する返信文書A-1を介して一の往信-返信関係で接続されている。こうした往信-返信関係で接続される全ての文書のまとまり(つまり、投稿文書A、返信文書A-1、返信文書A-1-a)がスレッドTHとなる。図3の系列図表には、図3に表されている限りにおいて、「投稿文書A、返信文書A-1、返信文書A-1-a」、「投稿文書A、返信文書A-1、返信文書A-1-b」、「投稿文書A、返信文書A-2」、「投稿文書A、返信文書A-3、返信文書A-3-a」、「投稿文書A、返信文書A-4」、「投稿文書B、返信文書B-1、返信文書B-1-a」、「投稿文書Z、返信文書Z-1、返信文書Z-1-a」という7個のスレッドTHが存在している。

【0041】データベースサーバ200のROM230内には、上記の系列図表を作成する処理(以下、系列図表作成処理という)を記述したプログラムが格納されている。このプログラムは、CPU220によって実行されることにより、多数の文書を文書群としてのスレッドごとに分類する分類手段として機能する。この系列図表作成処理の内容および処理手順を系列図表作成処理ルーチンとして図4に示す。本ルーチンは、投稿文書や返信

文書が書き込まれたときにCPU220が実行する処理である。本ルーチンが起動されると、まず、書き込まれた文書の宛先とされた文書を系列図表中で特定する処理を行なう(ステップS100)。次に、特定された文書の下層に、特定された文書と関連付けながら、書き込まれた文書の一部(例えば、表題等)を追加する処理を行なう(ステップS120)。次に、系列図表中の投稿文書や返信文書をスレッドごとに分類し直し(ステップS140)、分類後の系列図表をデータベース蓄積部260に更新して記憶して(ステップS160)、本ルーチンを終了する。

【0042】データベースサーバ200のROM230には、上記の系列図表作成処理の他に、電子掲示板に書き込まれた投稿文書および返信文書の要約を作成し、この作成された要約を系列図表と共にパーソナルコンピュータ900のディスプレイ上に表示する処理(以下、要約作成・表示処理という)の内容が記述されたプログラムが格納されている。この要約作成・表示処理の内容および処理手順を要約作成・表示処理ルーチンとして図5に示す。本ルーチンは、電子掲示板に書き込まれた文書の要約を作成する旨の実行指示がなされたときにCPU220が実行する処理である。

【0043】本ルーチンが起動されると、まず、①スレッド特定処理(ステップS200)を実行する。スレッド特定処理は、系列図表の中から要約を作成する対象となる一のスレッドTHを特定する処理である。以後、この特定された一のスレッドTHを対象として、②有効本文特定処理(ステップS210)、③重要文抽出処理(ステップS220)、④要約作成処理(ステップS240)、⑤表題作成処理(ステップS260)を実行する。これらの処理の詳細については後述する。これらのステップS220～S260の処理が全てのスレッドTHについて終了したときに(ステップS270)、後述する⑥表示データ出力処理(ステップS280)を実行して、本ルーチンを終了する。

【0044】以下、②有効本文特定処理、③重要文抽出処理、④要約作成処理、⑤表題作成処理、⑥表示データ出力処理の内容について順に説明する。なお、③重要文抽出処理には全体評価後切出処理が含まれており、④要約作成処理には文標準化処理が含まれているが、これらの処理についても③重要文抽出処理、④要約作成処理と併せて説明する。

【0045】②有効本文特定処理(図5のステップS210) 有効本文特定処理は、スレッド特定処理により特定されたスレッドTHにつき、該スレッドTH内に属する各文書の本文から文書の要旨を構成し得ない表現を抽出し、抽出された表現を除外した本文(以下、有効本文という)を特定する処理である。具体的には、図3で二点鎖線で囲ったスレッドTHの場合には、投稿文書Aの本文

を読み出し、挨拶文や投稿者の署名等の文書の主題の特定に関係の薄い表現を本文から除外し、削除後の本文を有効本文とする。返信文書A-1や返信文書A-1-aについても、本文を読み出して、投稿文書Aと同様の処理を行なう。但し、返信文書A-1や返信文書A-1-aの場合には、投稿文書Aの本文を引用した箇所についても本文から削除し、削除後の本文を有効本文とする。こうした有効本文は、RAM240の所定領域に一時的に記憶される。

【0046】図3で二点鎖線で囲ったスレッドTHに属する各文書(投稿文書A、返信文書A-1および返信文書A-1-a)の例文を図6に示す。投稿文書Aは甲により作成されて電子掲示板に書き込まれた文書であり、返信文書A-1は、投稿文書Aに対する返信として乙により作成され、電子掲示板に書き込まれた文書である。返信文書A-1-aは返信文書A-1に対する返信として甲により作成され、電子掲示板に書き込まれた文書である。図6に示すように、返信文書A-1、返信文書A-1、返信文書A-1-aには、それぞれ4つの文、6つの文、5つの文が含まれているが、上記の有効本文特定処理の実行により、返信文書A-1の第1文の「いつもお世話になります。」という定型的な挨拶文は削除され、残りの5つの文が有効本文とされる。なお、後の説明をわかりやすくするため、図6における各文の文頭に、有効本文とされる文の文番号(1～14)を示している。

【0047】③重要文抽出処理(図5のステップS220)

重要文抽出処理は、有効本文特定処理により特定された各文書の有効本文から各文書に所定の相関を有する文を重要文として抽出する処理である。本実施例では、重要文抽出処理において、一のスレッドTH内に属する全有効本文によって表わされる内容を考慮し、各文書に所定の相関を有する文として、全有効本文によって表わされる内容に最も近い意味の文を各文書の本文から切り出す処理(以下、全体評価後切出処理という)を行なう。この全体評価後切出処理の実行により切り出された文が重要文として抽出される。

【0048】重要文の抽出処理には、種々の手法が考えられるが、この実施例では、次の手法を採用した。重要文抽出処理の内容及び手順を、重要文抽出処理ルーチンとして図7に示す。本ルーチンは有効本文特定処理の終了後に起動する。本ルーチンが起動されると、まず、電子掲示板が有する各スレッドTHの種類を判別する処理を行なう(ステップS300)。スレッドTHの種類に応じてスレッドTHに含まれる各文書からの重要文の抽出手法を変えるためである。このスレッドの種類としては、例えば、スレッド内の各文書により最初の投稿文書の主題とは異なった主題に展開されているか否か(スレッド内の文書の現実の内容)、Q&A方式のスレッドか

否か（スレッドの形式）、スレッドがチャットのように往信と返信が頻繁に繰り返される性質のものか否か（スレッドの性質、特質）、スレッド内の一つの文書の長さが長文若しくは短文のいずれの傾向にあるか（スレッド内の文書の傾向）等を考えることができる。本実施例では、スレッドTHの種類を、質問型、複数話題型、会話型の3つのうちのいずれかに判別することとしている。

【0049】スレッドTHの種類判別は、以下の要領で行なうことができる。以下、図3で二点鎖線で囲ったスレッドTHを例として説明する。まず、主題を提供する最初の文書である投稿文書Aの有効本文を取り出し、有効本文内に「～の質問」や「～がわからない」等の表現がある場合には、このスレッドTHは質問型であると判別される。質問型にあてはまらない場合には、スレッドTH内の全有効本文（投稿文書A、返信文書A-1および返信文書A-1-aの有効本文）を取り出し、全有効本文の一定の範囲内に複数の主題があるか否かを判定する。複数の主題がある場合には、このスレッドTHは複数話題型であると判別され、複数の主題がない場合には、このスレッドTHは会話型であると判別される。なお、複数の主題の有無の判定は、全有効本文をTFIDF法を用いてベクトル表現に変換することにより行なうことができるが、ここでは詳しい説明は省略する。

【0050】こうしてスレッドTHの種類を判別した後、スレッドTHの種類に応じて重要文を抽出する処理を行なう（ステップS320）。この処理の詳細につき、以下、質問型、複数話題型、会話型に分けて説明する。

【0051】スレッドTHの種類が質問型である場合の重要文の抽出処理手法を図8のフローチャートに示す。質問型である場合には、まず、主題を提供する最初の文書である（図3の例では投稿文書A）の有効本文から「～の質問」や「～がわからない」等の表現を含む文を抽出し、これを最初の文書の重要文とする（ステップS400、S410）。次に、最初の文書と同じスレッドTH内に最初の文書と同じ者により作成された他の文書（図3の例では返信文書A-1-a）を探し、他の文書の有効本文から「うまくいった」や「解決した」、「わかりました」等の表現を含む文を抽出する。抽出された文が、最初の文書と同じ者により作成された他の文書の重要文となる（ステップS420、S430）。続いて、上記他の文書の有効本文の「うまくいった」等の表現を含む文の前後において質問に対する解決策が含まれている文を取り出し、この文に含まれる単語を解決策を示す単語として切り出す。次に、最初の文書とは異なる者に

$$TFIDF(d, t) = TF(d, t) \times Idf(t) \quad \dots (1)$$

但し：TF(d, t)は、各文d内において単語tが出現する回数、Idf(t)は、次式(2)による。

$$Idf(t) = \text{LOG}_e \{DB(db) / f(t, db)\} \quad \dots (2)$$

ここで、DB(db)は、全有効本文を構成する文の数、従って、図6に示した例では、値14、f(t, d

より作成された他人の文書（図3の例では返信文書A-1）の有効本文から上記解決策を示す単語が含まれている文を抽出する。抽出された文が、他人の文書の重要文となる（ステップS440～S460）。

【0052】このように、本実施例では、スレッドTHの種類が質問型である場合には、スレッドTH内の各文書から該文書に所定の相関を有する文（本実施例では、全有効本文によって表わされる内容に最も近い意味の文）を切り出す基準を、該文書群に属する文書の種類（図3の例では投稿文書A、返信文書A-1、返信文書A-1-a）に応じて異ならせている。

【0053】スレッドTHの種類が複数話題型および会話型である場合の重要文の抽出処理手法を図9のフローチャートに示す。複数話題型および会話型の場合には、まず、一のスレッドTH内に属する全文書の有効本文（以下、全有効本文という）を取り出す処理を行なう（ステップS500）。

【0054】本実施例では、図6に示した例文を有するスレッドTHを会話型とみなし、図9に示す重要文の抽出処理を適用している。図6に示す例文の場合には、投稿文書Aの有効本文（図6に示す文番号1～4の文）、返信文書A-1の有効本文（図6に示す文番号5～9の文）および返信文書A-1-aの有効本文（図6に示す文番号10～14の文）の計14個の文が、全有効本文として取り出される。

【0055】なお、複数話題型の場合には、ステップS500の処理において、全有効本文のうち、ある一の主題に属する本文と他の主題に属する本文とが別々に取り出され、取り出されたそれぞれの本文を全有効本文としてステップS510以下の処理を行なう。

【0056】次に、全有効本文から名詞である単語を抽出し、抽出された単語が全有効本文を構成する各文において隔って頻出する程度を求める処理を行なう（ステップS510）。

【0057】単語が隔って頻出する程度は、その単語が、全有効本文内で出現する回数を全有効本文を構成する文の数により正規化した値によって評価することができる。これは、例えばTFIDFとして知られている。TFIDFは、次の式で定義される。なお、以下の式で、dbは、対象となっているひとまとまりのテキストデータ（ここでは全有効本文）であり、dは、全有効本文を構成している各文、tはこのテキストに含まれる単語、とする。

【0058】

現する回数、Idfは、次式(2)による。

b)は、全有効本文において、単語tが出現する文の数、である。

【0059】図10は、図6に示した例文において、該例文に含まれている名詞である単語の出現頻度を数えた結果を示す。こうして名詞である単語の出現頻度を求めた後、本実施例では、出現頻度が2以上の単語につき、上式に基づいてTFIDF値を求めた。こうして求められた各単語のTFIDF値を図11に示す。

【0060】続いて、各文を構成する名詞である単語のTFIDFを合算する処理を行なう(ステップS520)。この合算処理の結果を図11の合算値の欄に示した。

【0061】次に、スレッドTHに含まれる各文書の各有効本文ごとに各文の合算値の大小を判定し、合算値が最も大きい2つの文を抽出する処理を行なう(ステップS530)。図11に星印付きで示すように、図6の例文の場合には、投稿文書Aからは文番号1、2の文が、返信文書A-1からは文番号7、8の文が、および返信文書A-1-aからは文番号11、14の文が、それぞれ抽出される。

【0062】以上のステップS510～S530までの処理が、前述した全体評価後抽出処理に相当する。この全体評価後抽出処理によれば、各有効本文内における出現頻度が高くなければ、単語のTFIDF値は高くない。反面、全有効本文にまんべんなく出現するような単語(例えば、「こと」や「場合」)の場合には、f(t, d)が大きな値となるために、TFIDF値は結局小さな値となる。つまり、各有効本文を構成する各文において高い頻度で出現する単語が存在する文ほど、TFIDF値は大きな値となり、しかもそういう単語が多いほど、合算値は高い値となるのである。

【0063】以上のように、スレッドTHに含まれる各文書から2つの重要文が抽出される。この後、抽出された重要文をRAM240の所定領域に記憶して(図7のステップS340)本ルーチンを終了し、要約作成処理(図5のステップS240)に移る。

【0064】㊸要約作成処理(図5のステップS240)

要約作成処理次に、重要文抽出処理により抽出された重要文に基づいてスレッドTHの要約を作成する処理である。図3の例では、投稿文書A、返信文書A-1および返信文書A-1-aからなるスレッドTHの要約が作成される。この要約の作成は、具体的には、スレッド内の各文書から抽出された重要文を所定のアルゴリズムに基づいて組み合わせることにより作成される。例えば、図3に二点鎖線で囲ったスレッドTHが質問型であり、投稿文書Aが質問を含み、返信文書A-1が解決策の提案を含み、返信文書A-1-aが解決したという結果を含む場合には、投稿文書Aから抽出された重要文と返信文書A-1-aから抽出された重要文とを組み合わせるウエイトを大きくして、要約を作成すればよい。「…という疑問点が解決した」という文は、一般に、閲覧者に興

味を生じさせると考えられるからである。

【0065】本実施例では、重要文に重要文抽出処理により抽出された重要文を標準的な表現に書き換える処理(以下、文標準化処理という)を行なうことにより、要約を作成している。具体的には、文標準化処理において、冗長な表現の短文化や自立語の他の自立語への置換、付属語の用法の訂正、方言から標準語への置換、予め定めた文字への置き換え、表記のゆれの統一等の処理を行なう。標準化の処理には大別すると、文字の標準化、表記のゆれの統一、自立語の統一などを考えることができる。これらの処理については説明は省略するが、大まかな例示を挙げると以下の通りである。

【0066】文字の標準化を例示すると、

- (A) 括弧:『』と「」の置き換えを行なうなど、
- (B) 引用符:“”と””の置き換えを行なうなど、
- (C) 一般記号:種々の記号(例えば「:」,「?」など)について、半角/全角の置き換えを行なうなど、
- (D) カタカナや英数字:全角/半角や大文字/小文字の置き換えを行なうなど、
- (E) 句点・読点:句点、読点を「,」「。」に統一するなど、

(F) 名前の贅文字:「クイーン=エリザベス」を「クイーン・エリザベス」に置き換えるなど、がある。

【0067】表記のゆれとは、日本語における表記の曖昧さ、許容幅を言い、例えば、

- ㊶長音記号のゆれ:例、ウィンドウズ、ウィンドーズ、
- ㊷送り仮名のゆれ:例、売上げ、売り上げ、
- ㊸拗音表記のゆれ:例、ウィザード、ワイザード、
- ㊹複合語のかな表記のゆれ:例、売り上げ、売りあげ、
- ㊺外来語表記のゆれ:エンゼル、エンジェル、
- ㊻繰り返し文字のゆれ:例、正正堂堂、正々堂々、などを挙げることができる。

【0068】更に、自立語の統一処理としては、

- (イ) 修飾語:すごく、最高の、一番の、高い、など
- (ロ) 名詞:パーソナルコンピュータ、パソコン、PC、など
- (ハ) 動詞:知らせる、連絡する、通知する、などを例示することができる。

【0069】こうした処理を行なうための基準データやルールは、データベースサーバ200のハードディスク270内の辞書に格納されている。予め定めた文字への置き換えを例にとると、「PC」といった単語は「パソコン」という単語に置き換えられる。これにより、一のスレッドTH内における文書や重要文の用語を統一することができる。

【0070】㊼表題作成処理(図5のステップS260)

表題作成処理は、重要文抽出処理により抽出された重要文に基づいてスレッドの表題を作成する処理である。図3の例では、投稿文書A、返信文書A-1および返信文

書A-1-aからなるスレッドTHの表題が作成される。具体的には、重要文の文法情報を解析し、必要な単語を切り出して付属語を補完することにより、表題を作成する。

【0071】⑤表示データ出力処理(図5のステップS280)

表示データ出力処理は、上記作成された要約、表題および系列図表を表示データとして出力する処理である。この処理は、図5のステップS270の処理において電子掲示板の全てのスレッドTHについて要約および表題が作成されたと判断されたときに実行される。

【0072】表示データ出力処理において出力されるデータイメージの一例を図12に示す。図12では、図3に示した系列図表に基づくデータイメージを表している。このデータイメージは、表示データがデータベースサーバ200からネットワーク100を介してクライアントのパーソナルコンピュータ900に到達したときに、パーソナルコンピュータ900のディスプレイのプレビュー画面に初期表示される内容である。

【0073】図12に示すように、プレビュー画面には、各スレッドごとに要約および表題が表示されている。スレッド名は、系列図表の最下層に位置する返信文書に基づいて付けられており、例えば、図3で二点鎖線で囲ったスレッドTHのスレッド名は、最下層の返信文書名である「A-1-a」とされている。なお、本実施例では、投稿日(投稿文書が書き込まれた日)、更新日(返信文書が書き込まれた最近の日)、階層数(スレッド内における返信文書の数)が併せて表示されるので、これらの表示を話題の選択に関する参考資料とすることができる。

【0074】各スレッドについての表題および要約は時系列順に一覧表示される。本実施例では、更新日が新しいスレッドから順に表示している。図12で言えば、スレッド名が「B-1-a」のものが最も最近に更新されたスレッドである。これにより、最新の主題や話題に容易にアクセスすることができる。

【0075】なお、図12に示すプレビュー画面において、表題や要約等の他、各スレッドに属する投稿文書や返信文書の作成者の名前を表示することとしてもよい。こうすれば、各作成者による発言内容の変化を把握しやすくなる。

【0076】クライアントは、プレビュー画面において表題や要約を参照し、興味のある話題のスレッドを選択することができる。こうしてスレッドが選択されると、選択されたスレッド内に属する全ての文書情報が、文書の往信-返信関係と共に、一覧表示画面に表示される。スレッド名が「A-1-a」のスレッドTHが選択されたときに一覧表示画面に表示される内容を図13に示す。この一覧表示画面により、興味のある話題の内容を詳細に確認することができる。

【0077】以上説明した本実施例の掲示板表示システムKSは、往信-返信関係を有する投稿文書、投稿に対する返信文書、返信に対する返信文書から構成されたスレッドTHに関し、スレッドTHに属する文書全体によって表わされる内容を考慮して、スレッドTHに属する各文書から文書全体の内容に最も近い意味の文を重要文として抽出し、抽出された重要文に基づいてスレッドTHの要約を作成し、作成された要約を画面上に表示する。従って、スレッドTHごとの話題の内容を迅速かつ的確に把握することが可能となり、話題の取捨選択を迅速かつ容易に行なうことができる。

【0078】本実施例の掲示板表示システムKSは、スレッドTHの種類を判別し、判別されたスレッドTHの種類に応じて重要文を抽出し、この重要文を表示する。従って、主題の展開や話題の内容をより正確に把握することができる。

【0079】本実施例の掲示板表示システムKSは、抽出された重要文に文標準化手段を適用して、重要文を標準的な表現に書き換え、要約を作成する。こうすれば、文書作成者の文書作成に関する技量に拘らず均一なレベルの文がスレッドの要旨として表示されるので、文書群ごとの話題の内容をより一層把握しやすくなる。例えば、冗長な表現を有する重要文を短文化した場合には、限られた表示領域において内容の把握しやすい文を表示することができる。

【0080】本実施例の掲示板表示システムKSは、各スレッドTHに属する文書から文書の要旨を構成し得ない表現を除外した有効本文に基づいて重要文の抽出または要約の作成を行なう。従って、重要文の抽出または要約の作成を効率的かつ精度よく行なうことができる。

【0081】本実施例の掲示板表示システムKSは、重要文の抽出結果を利用してスレッドTHの表題を作成し、スレッドTHの要約と共に表示する。従って、作成者によって付けられた各文書の表題を羅列する場合と比較して、各スレッドTHの話題の種類を把握しやすくなる。

【0082】(2)第2実施例：往信-返信関係を有する文書には、電子掲示板への書き込み文書以外に、発信者と特定の受領者との間で文書を送り合う電子メールがある。このような電子メールのまとまりに対して重要文を抽出し、要約を作成する実施例をメール表示システムMSとして説明する。図14は、本発明の第2実施例であるメール表示システムMSの概略構成を示す説明図である。図14に示すように、メール表示システムMSは、掲示板表示システムKSとほぼ共通の構成を備えているが、データベースサーバ200を備えない点で掲示板表示システムKSと異なる。なお、図14では、この共通の構成につき、符号の上二桁を図1と同じ数字を用いて表わしている。

【0083】配信サーバ30、31のROMには、特定

の発信者と特定の受領者との間で送受信されたメールのまとまり（以下、メール束という）を受信ないし送信日時順に整理して時系列図表を作成する処理を記述したプログラムが格納されている。このプログラムが実行されると、図15に示すような時系列図表が作成される。図15に示すように、X氏からY氏に送信したメールX1を受け取ったY氏は、X氏に対してメールY1を送信し、このメールY1を受け取ったX氏は、Y氏に対してメールX2を送信している。図15では、こうしたX氏とY氏との間のメールの送受信が5回続いている。これにより、X氏とY氏との間でやりとりされたメールは、「メールX1→メールY1→メールX2→メールY1→メールX3→メールY3→メールX4→メールY4→メールX5→メールY5」というように時系列順に連鎖している。このように連鎖する合計10通のメールがメール束MBとなる。なお、時系列図表中の各メールの本文は、時系列図表とリンクされて配信サーバ30、31内のハードディスクに蓄積されている。

【0084】配信サーバ30、31のROMには、上記の第1実施例と同様の、要約作成・表示処理の内容が記述されたプログラムが格納されている。このプログラムが配信サーバ30、31のCPUの命令によって実行されることにより、メール束に対して、第1実施例におけるスレッドTHと同様に、①スレッド特定処理、有効本文特定処理、②重要文抽出処理、③要約作成処理、④表題作成処理、⑤表示データ出力処理と同様の処理が行われる。この結果、メール束に属するメール全体によって表わされる内容を考慮して、メール束に属する各メールからメール全体の内容に最も近い意味の文が重要文として抽出され、抽出された重要文に基づいてメール束の要約が作成される。こうしたメール束の要約は、X氏のパーソナルコンピュータ90とY氏のパーソナルコンピュータ91の双方のレビュー画面上に表示される。従って、クライアントは、メール束ごとの話題の内容を迅速かつ的確に把握することが可能となり、メールの取捨選択を迅速かつ容易に行なうことができる。

【0085】なお、上記の実施例では、メールの本文の表示を例として取り上げたが、こうした表示の手法は、メール本文の表示に限定されるものではなく、複数の文から構成されたデータについても適用することができる。例えば、論文データベースの抄録文を対象として、抄録文の中から更に重要文を抽出し、これを論文の表題と共に表示するといった態様で実現することができる。あるいは電子カルテの記載や、電子的に配信される新聞などの記事から重要文を抽出して表示するといった実現形態も存在する。抽出した重要文は、モニタに表示するだけでなく、モニタへの表示に代えて、あるいは表示と共に、音声により読み上げることも好適である。もとより、プリンタなどで印字することも差し支えない。

【0086】以上、本発明の実施の形態について説明し

たが、本発明はこうした実施の形態に何等限定されるものではなく、例えば、重要文を他とは異なる態様（例えば色違い）で表示するなど、本発明の要旨を逸脱しない範囲内において、更に種々なる形態で実施し得ることは勿論である。

【0087】第1実施例では、系列図表作成処理（図4）や要約作成・表示処理（図5）をデータベースサーバ200が実行することとしたが、系列図表作成処理や要約作成・表示処理の内容を記述したプログラムをデータベースサーバ200以外の装置に格納し、データベースサーバ200以外の装置が実行する構成としてもよい。例えば、配信サーバ300、310、320やクライアントのパーソナルコンピュータ900、910、920が実行する構成等を考えることができる。

【0088】また、第2実施例では、系列図表作成処理や要約作成・表示処理を配信サーバ30、31が実行することとしたが、系列図表作成処理や要約作成・表示処理の内容を記述したプログラムを配信サーバ30、31以外の装置に格納し、配信サーバ30、31以外の装置が実行する構成としてもよい。例えば、上記プログラムを読み取ったパーソナルコンピュータ90、91が実行する構成等を考えることができる。

【0089】また、上記実施例では、重要文抽出処理において、電子掲示板への書き込み文書が、平均的な文の数が約5つ程度であり、一文が比較的短い傾向にあることを考慮し、抽出する文の数を2つと定めたが（図9のステップS530の処理）、抽出する文の数は、スレッドTHに含まれる各文書（図3で言えば、投稿文書A、返信文書A-1、返信文書A-1-a）の長さ等に応じて任意に定めることができる。

【0090】上記実施例では、重要文を抽出した後、この重要文に対して文標準化処理を行なったが、重要文を抽出する前に有効本文（投稿文書や返信文書の本文等）に対して文標準化処理を行なう構成としてもよい。

【0091】上記実施例では、重要文抽出処理（図5のステップS220）の終了後、要約作成処理（図5のステップS240）、表題作成処理（図5のステップS260）を行なったが、この要約作成処理や表題作成処理を行なうことなく、重要文抽出処理で抽出された各文書の重要文を表示データ出力処理（図5のステップS280）により表示可能に出力する構成としてもよい。この構成によっても、スレッドTHメール束に属する文書全体によって表わされる内容を考慮して、スレッドTHに属する各文書から文書全体の内容に最も近い意味の文が重要文として抽出され、抽出された重要文が画面上に表示される。従って、各スレッドTHや各メール束における主題の展開を、表示された重要文を見ただけで正確に把握することができる。

【0092】特に、メール束に属する各メールから重要文を抽出し、この重要文を時系列に配列した場合には、

このように配列された重要文をモニタ上に表示したり、プリントアウトすることにより、過去のやり取りを一目で把握することが可能となり、ビジネスシーン等において有効に活用することができる。過去のやり取りにおける特定の内容（例えば、双方が合意した納期）を調べようとする際に、メール束に属する各メールを順次にモニタに表示したり若しくはプリントアウトし、メール束から特定の内容を探すという必要がないからである。

【0093】以上本発明のいくつかの実施例と変形例について説明したが、本発明は、これらの実施例に何ら限定されるものではなく、本発明の要旨を逸脱しない範囲内において種々の態様で実施することができる。例えば、複数の文書の関係は、往信・返信の関係に限られるものではなく、例えば、同一の作成者が作成した複数の文書、コメントや日記、報告書などであっても差し支えない。また、本発明の文書表示を行なう部分は、複数の文書を蓄積しているサーバ上に置いても良いし、複数の文書を閲覧するクライアント側においても良い。あるいはこれらのサーバ・クライアントシステムの中間に、専用のサーバとして置くことも可能である。これらのシステムを構成するためのプログラムは、CD-ROMなどの媒体の形態で取り扱っても良いし、プログラムをサーバ上に置き、使用するものがこれをダウンロードして、自らのコンピュータ上で、上述した文書表示システムの少なくとも一部を構築するものとしても良い。

【図面の簡単な説明】

【図1】本発明の第1実施例である掲示板表示システムKSの概略構成を示す説明図である。

【図2】データベースサーバ200の構造を示す説明図である。

【図3】電子掲示板に書き込まれている文書の系列図表を示す説明図である。

【図4】系列図表作成処理ルーチンを示すフローチャートである。

【図5】要約作成・表示処理ルーチンを示すフローチャートである。

【図6】図3で二点鎖線で囲ったスレッドTHに属する各文書の例文を示す説明図である。

【図7】重要文抽出処理ルーチンを示すフローチャート

である。

【図8】スレッドTHの種類が質問型である場合の重要文の抽出処理を示すフローチャートである。

【図9】スレッドTHの種類が複数話題型および会話型である場合の重要文の抽出処理を示すフローチャートである。

【図10】図9の例文を用いて各単語のTFIDF値を計算した計算例を示す説明図である。

【図11】図9の例文におけるTFIDF値の文単位の合算値を例示する説明図である。

【図12】表示データの出力後にパーソナルコンピュータ900のディスプレイのプレビュー画面に初期表示される内容を示す説明図である。

【図13】一覧表示画面の一例を示す説明図である。

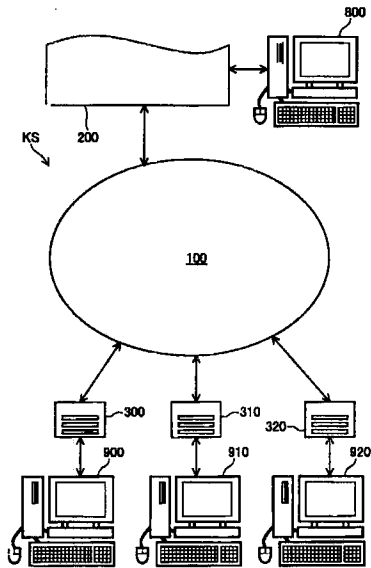
【図14】メール表示システムMSの概略構成を他の実施例として示す説明図である。

【図15】配信サーバ30、31により作成されるメールの時系列図表を示す説明図である。

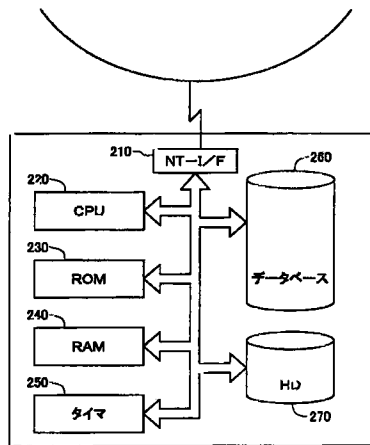
【符号の説明】

10…ネットワーク
30、31…配信サーバ
90、91…パーソナルコンピュータ
100…ネットワーク
200…データベースサーバ
210…ネットワークインタフェース
220…CPU
230…ROM
240…RAM
250…タイマ
260…データベース蓄積部
270…ハードディスク
300、310、320…配信サーバ
800…パーソナルコンピュータ
900、910、920…パーソナルコンピュータ
KS…掲示板表示システム
MS…メール表示システム
MB…メール束
TH…スレッド

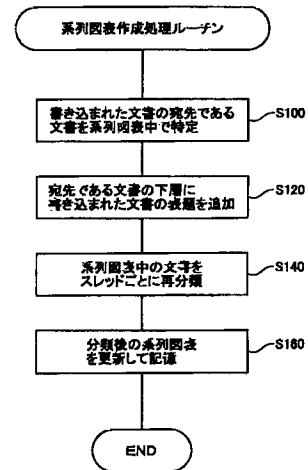
【図1】



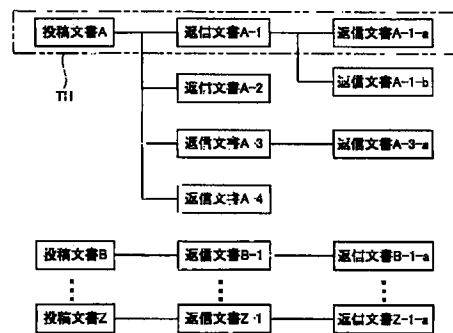
【図2】



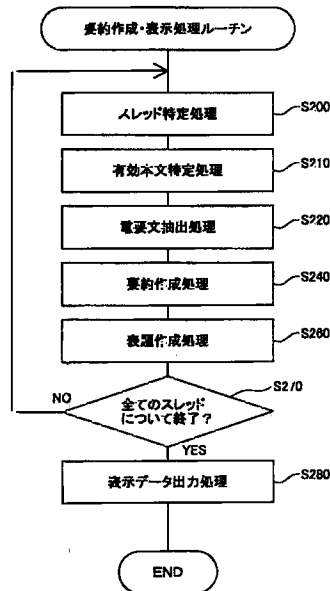
【図4】



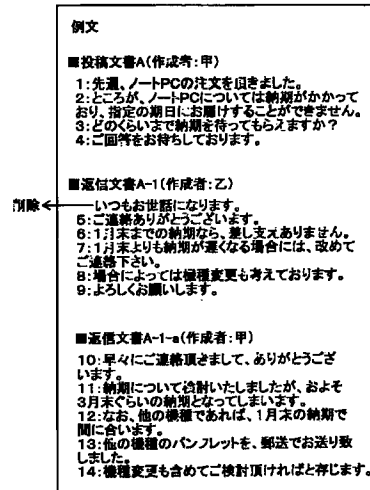
【図3】



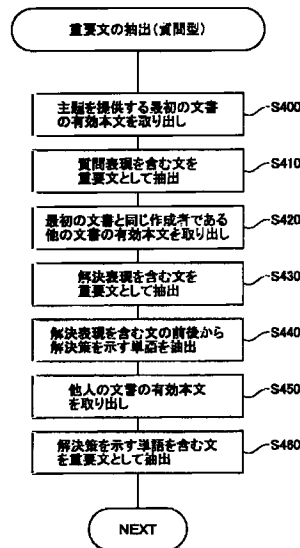
【図5】



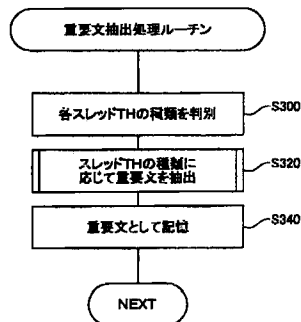
【図6】



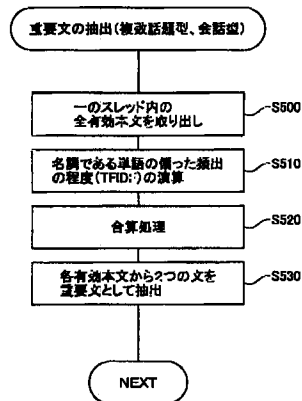
【図8】



【図7】



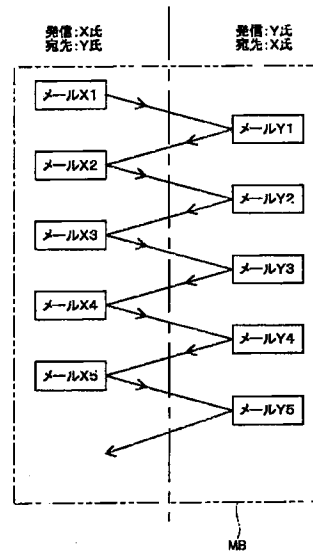
【図9】



【図10】

文番号	先通	ノート	パソコン	注文	納期	指定	期日	団体	連絡	機種	変更	パンフレット
1	1	1	1	1								
2	1	1	1	1	1	1	1					
3					1							
4								1				
5									1			
6				1								
7				1					1			
8										1	1	
9												
10									1			
11				2								
12										1		
13										1		1
14										1	1	
合計	1	2	2	1	5	1	1	1	3	4	2	1

【図15】



【図11】

出現頻度2以上の単語についての演算結果

文番号	合算値	ノート	パソコン	納期	連絡	機種	変更
★ 1	3.892	1.946	1.946				
★ 2	4.739	1.946	1.946	0.847			
3	0.847			0.847			
4	1.946						
5	1.540				1.540		
6	0.847			0.847			
★ 7	2.387			0.847	1.540		
★ 8	3.199					1.253	1.946
9	0						
10	1.540				1.540		
★ 11	1.885			1.885			
12	1.253					1.253	
13	1.253					1.253	
★ 14	3.199					1.253	1.946

(巻6) 103-108278 (P2003-108278A)

【図12】

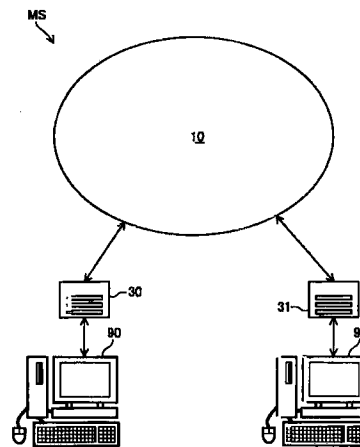
スレッド名	表題	要約	投稿日	更新日	層数
B-1-a	11/28	12/5	3
A-1-a	ノートパソコンの 納期について	ノートパソコンは納期がかかっており、 機種の変更が必要な場合がある。	11/25	11/30	3
A-4	11/25	11/28	2
A-1-b	11/25	11/27	3

【図13】

スレッド名: A-1-a	
作成者: 甲	書込日: 2001/11/25
<p>先週、ノートPCの注文を頂きました。ところが、ノートPCについては納期がかかっており、指定の期日にお届けすることができません。どのくらいまで納期を待っていただけますか？ご回答をお待ちしております。</p>	
作成者: 乙	書込日: 2001/11/28
<p>いつもお世話になります。ご連絡ありがとうございます。 1月末までの納期なら、差し支えありません。 1月末よりも納期が遅くなる場合には、改めてご連絡下さい。 場合によっては機種変更も考えております。よろしくお願ひします。</p>	
作成者: 甲	書込日: 2001/11/30
<p>早々にご連絡頂きまして、ありがとうございます。納期について検討いたしましたが、おおよそ3月末ぐらいの納期となってしまいます。 なお、他の機種であれば、1月末の納期で間に合います。 他の機種の手配を、郵送でお送り致しました。 機種変更も含めてご検討頂ければと存じます。</p>	

(特 7) 103-108278 (P2003-108278A)

【図14】



フロントページの続き

(72)発明者 長保 和義
長野県諏訪市大和三丁目3番5号 セイコ
ーエプソン株式会社内

(72)発明者 鈴木 勝仁
長野県諏訪市大和三丁目3番5号 セイコ
ーエプソン株式会社内

Fターム(参考) 5E501 AA02 AB19 AC25 AC34 BA05
CA02 CB02 CB09 EA12 EB17
FA13 FA23 FB45